

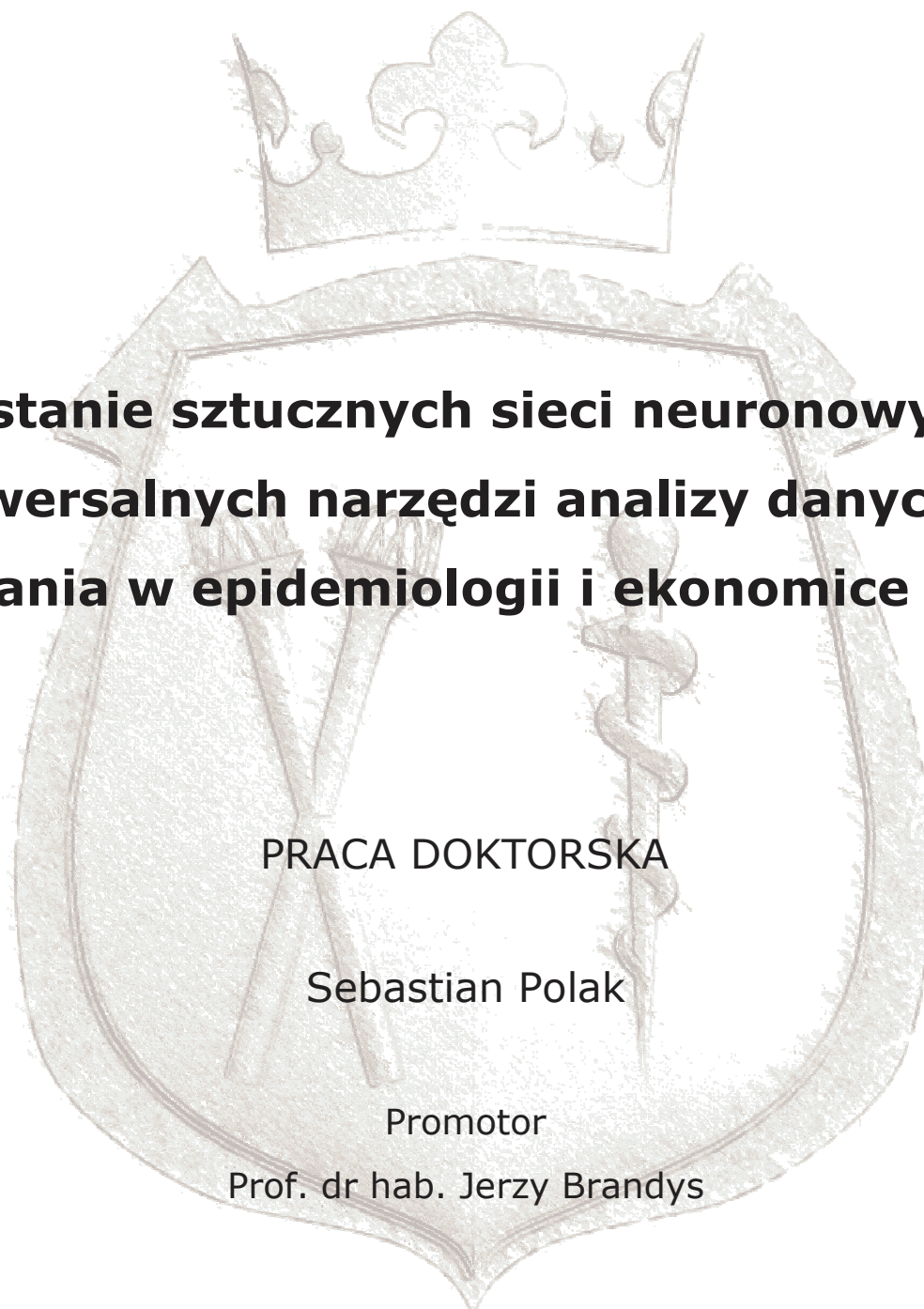
Pracownia Farmakoepidemiologii i Farmakoekonomiki

Katedra Toksykologii

Wydziału Farmaceutycznego

Collegium Medicum

Uniwersytetu Jagiellońskiego



**Wykorzystanie sztucznych sieci neuronowych jako  
uniwersalnych narzędzi analizy danych i  
modelowania w epidemiologii i ekonomice zdrowia**

PRACA DOKTORSKA

Sebastian Polak

Promotor

Prof. dr hab. Jerzy Brandys

Kraków 2005

1	Wstęp .....	3
1.1	Współczesne spojrzenie na zagadnienia magazynowania i analizy informacji .....	3
1.2	Modelowanie w naukach o zdrowiu.....	4
1.3	Ekonomizacja opieki medycznej.....	7
2	Podstawy teoretyczne .....	8
2.1	Sztuczne sieci neuronowe .....	8
2.1.1	Wstęp .....	8
2.1.2	Neurobiologiczne inspiracje dla obliczeń neuronowych.....	10
2.1.3	Podstawowy model sztucznego neuronu i sieci neuronowej..	10
2.1.4	Uczenie sieci neuronowych .....	16
2.1.5	Sieci neuronowe bazujące na logice rozmytej.....	33
2.1.6	Sieci neuronowe wykorzystujące algorytmy genetyczne .....	36
2.1.7	Przygotowanie danych uczących i testowych oraz dobór architektury sieci i współczynników uczenia .....	36
2.1.8	Analiza nauczonego modelu neuronowego .....	38
2.2	Farmakoekonomika.....	39
2.2.1	Wstęp .....	39
2.2.2	Etapy badania farmakoekonomicznego.....	40
2.2.3	Źródła danych wykorzystywanych w analizach farmakoekonomicznych.....	41
2.2.4	Typy analiz farmakoekonomicznych .....	43
2.2.5	Modelowanie w farmakoekonomice .....	64
2.3	Epidemiologia.....	66
2.3.1	Wstęp .....	66
2.3.2	Modelowanie w epidemiologii .....	67
2.4	Regresja logistyczna .....	75
3	Cel pracy .....	78
4	Oprogramowanie i sprzęt obliczeniowy .....	79
4.1	Oprogramowanie .....	79
4.2	Sprzęt obliczeniowy .....	80
5	Zastosowanie sztucznych sieci neuronowych w farmakoekonomice ..	81
5.1	Metodyka .....	81
5.1.1	Dane .....	81
5.1.2	Wartość wyjściowa (modelowana) .....	85
5.1.3	Ocena błędu systemów neuronowych.....	85
5.1.4	Walidacja systemów neuronowych.....	86
5.1.5	Schemat opisu testowanych architektur sieci neuronowych..	86
5.1.6	Formatowanie (preprocessing) danych .....	86
5.1.7	Przebieg pracy .....	91
5.2	Rezultaty poszczególnych etapów pracy .....	93
5.2.1	Modelowanie z wykorzystaniem danych pierwotnych – wyjście ciągle.....	93
5.2.2	Modelowanie z wykorzystaniem danych przeformatowanych	96
6	Zastosowanie sztucznych sieci neuronowych w epidemiologii .....	126
6.1	Metodyka .....	126

6.1.1	Dane .....	126
6.1.2.	Wartości wyjściowe (modelowane) .....	127
6.1.3.	Preprocessing i zaszumianie danych.....	128
6.2	Rezultaty poszczególnych etapów pracy – modelowanie ryzyka wystąpienia cukrzycy ciężarnych (BRFSS_1999) .....	129
6.2.1.	Opis problemu .....	129
6.2.2.	Przygotowanie danych .....	130
6.2.3.	Przebieg prac.....	132
6.3	Rezultaty poszczególnych etapów pracy – modelowanie ryzyka wystąpienia nadciśnienia tętniczego (BRFSS_1999) .....	136
6.3.1	Opis problemu .....	136
6.3.2	Przygotowanie danych .....	137
6.3.3.	Przebieg prac.....	140
6.4	Rezultaty poszczególnych etapów pracy – modelowanie ryzyka wystąpienia ostrego incydentu wieńcowego (NHEFS I) .....	146
6.4.1.	Opis problemu .....	146
6.4.2.	Przygotowanie danych .....	148
6.4.3.	Przebieg prac.....	150
6.5	Rezultaty poszczególnych etapów pracy – modelowanie ryzyka wystąpienia choroby nowotworowej (NHEFS I) .....	155
6.5.1.	Opis problemu .....	155
6.5.2.	Przygotowanie danych .....	155
6.5.3.	Przebieg prac.....	156
7	Dyskusja wyników i zastosowania praktyczne.....	161
7.1	Zastosowanie sztucznych sieci neuronowych w farmakoekonomice 161	
7.2	Zastosowanie sztucznych sieci neuronowych w epidemiologii ...	181
8	Wnioski .....	191
9	Streszczenie .....	192
10	Piśmiennictwo .....	194
11	Spis równań.....	208
12	Spis schematów .....	209
13	Spis wykresów .....	213
14	Spis tabel .....	214
15	Załączniki .....	219

# 1 Wstęp

## 1.1 Współczesne spojrzenie na zagadnienia magazynowania i analizy informacji

Jedną z cech charakteryzujących współczesny świat jest natłok informacji, podlegających procesowi magazynowania oraz wymagających analizy. Dane tworzone w trakcie prac naukowych, badawczych, doświadczeń oraz w codziennym życiu zawodowym zajmują zbyt duże zasoby aby móc gromadzić je w postaci tradycyjnej. Jest to jedna z przyczyn gwałtownego rozwoju i przyspieszonej ewolucji technologii cyfrowych, które są niewspółmiernie wydajniejsze niż podejście klasyczne. Terabajty informacji są przechowywane na nośnikach gwarantujących nie tylko większą trwałość ale i umożliwiających szybki dostęp do poszukiwanego fragmentu informacji. Dostęp do zasobów akumulowanej przez ludzkość wiedzy oraz umiejętność jej analizy i wyciągania na jej podstawie wniosków stanowi obecnie najcenniejszy kapitał.

W powyższe stwierdzenia doskonale wpasowuje się świat szeroko pojętej ochrony zdrowia wraz z szeroką gamą zagadnień pozostających w polu jej zainteresowań. Coraz węższe specjalizacje i postęp w dziedzinie farmacji, medycyny, epidemiologii ludności oraz innych dyscyplin jedynie potęguje zapotrzebowanie na specjalistyczną i wiarygodną informację.

Sam fakt zbierania i udostępniania informacji nie stanowi jeszcze o dokonującym się procesie rozwoju sposobów analizy danych, jest jednak jego pośrednią przyczyną. Powodem bezpośrednim jest konieczność dostarczenia specjalistom-praktykom opracowanego, przeanalizowanego i zwalidowanego quantum informacji możliwej do wykorzystania. Jest to więc zobrazowanie drugiej części pierwszego zdania pracy – potrzeby analizy danych. W niektórych sytuacjach informacje surowe, nie poddane procesowi obróbki są niezwykle trudne do bezpośredniego wykorzystania lub nawet wstępnej oceny. Nawet w przypadku korzystania z narzędzi

komputerowych i zaawansowanych rozwiązań programowych szybkie wyciągnięcie wniosków na podstawie bazy danych zbudowanej z setek tysięcy rekordów, z których każdy opisywany jest setkami parametrów nie jest możliwe. Powyższe stwierdzenie oraz ograniczenia klasycznej, opartej na konwencjonalnej statystyce analizy danych były przyczyną rozwoju nowatorskich, opartych m.in. o sztuczną inteligencję i logikę rozmytą dróg umożliwiających wnioskowanie na podstawie zgromadzonych doświadczeń. Jednym z narzędzi, które wykorzystywane są w analizie danych (w tym danych medycznych) są sztuczne sieci neuronowe.

## 1.2 Modelowanie w naukach o zdrowiu

Model to uproszczony schemat rzeczywistości, którą stara się w sposób syntetyczny opisać. Cechą charakterystyczną modelu jest wykorzystywanie najważniejszych cech przedstawianych sytuacji lub obiektów, w celu ich najlepszego scharakteryzowania. Celem tworzenia modeli w naukach o zdrowiu jest możliwość zastąpienia, odpowiednio przetestowanymi schematami często kosztownych lub niebezpiecznych dla poddawanych badaniom ludzi lub zwierząt eksperymentów. Przykładem jednego z najczęściej wykorzystywanych modeli i metod modelowania są badania kliniczne substancji chemicznych, o potencjalnym pozytywnym działaniu farmakologicznym przed zarejestrowaniem jako lek. Dzięki odpowiedniej procedurze oraz stworzeniu pewnej uproszczonej wizji rzeczywistości (kolejno na ochotnikach zdrowych oraz chorych), unika się konieczności przeprowadzania badań obejmujących duże populacje, a uzyskane na podstawie badań klinicznych wyniki i wnioski ekstrapoluje się na całe populacje chorych, będących potencjalną grupą docelową dla nowego leku.

Prawidłowo skonstruowany model umożliwia zarówno interpolację jak i ekstrapolowanie, poza zakres danych wykorzystywanych w trakcie jego budowania. Rolą modelowania jest również kojarzenie danych i dowodów pochodzących z różnych źródeł (stąd za jedną z uproszczonych metod modelowania uznawana jest metaanaliza) czy uogólnianie wyników.

Definicja modelowania nie określa narzędzia matematycznego wykorzystywanego do tworzenia modeli. Wśród możliwych instrumentów wymienić można m.in. statystyczne techniki regresyjne (liniowa, logistyczna, Bayesowska), narzędzia z zakresu inteligencji obliczeniowej (sztuczne sieci neuronowe, systemy bazujące na algorytmach genetycznych, sieci rozmyte), techniki matematyczne (drzewa decyzyjne). Wybór techniki modelowania zależy zarówno od charakterystyki rozwiązywanego problemu jak i ilości oraz jakości danych, wykorzystywanych w trakcie budowania modelu.

Modelowanie matematyczne lub numeryczne wykorzystuje szeroko rozumiane metody obliczeniowe ułatwiające budowanie systemów wspomagających podejmowanie decyzji. Przy pomocy formuł statystycznych i matematycznych opisywane są w uproszczeniu - lecz jednak pozwalający wyciągać pewne wnioski - sposoby zależności z zakresu tak różnych dziedzin jak fizyka jądrowa, meteorologia czy makroekonomia [18]. W farmacji i medycynie opisywane systemy są wykorzystywane zarówno w farmakokinetyce i farmakologii z jednej strony jak i farmakoekonomicznie oraz ogólnej ocenie technologii medycznych [135]. Przykłady modelowania w opiece zdrowotnej to m.in.:

- modele przeżycia 30-to dniowego i rocznego pacjentów po ciężkim zawale serca [128]
- przewidywanie wykorzystania zasobów wykorzystywanych na oddziałach szpitalnych [115]
- budowa modelu decyzyjnego postępowania w infekcjach cytomegalowirusem (CMV) obejmującego zarówno aspekty zdrowotne jak i ekonomiczne [8]
- analiza czynników ryzyka syndromu BOS (*Bronchiolitis Obliterans Syndrome*) po transplantacji płuc [42]
- analiza przeżycia pacjentek z nowotworem piersi [109]
- modelowanie badań farmakoekonomicznych przy użyciu techniki szeregów czasowych (*Temporal Objects*) [50]

Przegląd literatury z zakresu modelowania w naukach o zdrowiu wskazuje na obszary, które są szczególnie predysponowane do wykorzystywania modeli. Wśród wielu innych zwraca uwagę rosnąca rola modelowania w analizie przeżycia [93] oraz w farmakoekonomice [95].

Zgodnie z międzynarodowymi normami modele muszą spełniać kryteria jakościowe i merytoryczne, wymienione w wytycznych dotyczących dobrej praktyki modelowania w naukach medycznych [135]. Czynniki decydujące o jakości modelu oraz podlegające ocenie w procesie walidacji to [135]:

- struktura modelu – powinna odpowiadać problemowi, który jest rozwiązywany
- dane wykorzystane do budowy modelu – ocenie podlegają zarówno źródła danych (badania kliniczne, epidemiologiczne) jak i ich zgodność z rozwiązywanym problemem (brak przypadkowości)
- zgodność wewnętrzna i zewnętrzna – ich zapewnienie wymaga wprowadzenia procedury kontroli procesu modelowania w trakcie budowania modelu (na poziomie wprowadzania danych, kalkulacji modelu) jak i po jego zbudowaniu i zastosowaniu (porównanie z dostępnymi wynikami innych badań naukowych)

Skuteczność nie zawsze stanowi najważniejszy wskaźnik oceny modelu. Niemniej jednak w specyficznych zastosowaniach, szczególnie w przypadku modeli predykcyjnych, jest to ważny fragment oceny.

Obecnie coraz szerzej stosowane są metody modelowania oparte na technikach obliczeniowych. Wśród wielu technik modelowania zaliczanych do wspomnianej grupy wymienić można sztuczne sieci neuronowe czy układy wykorzystujące algorytmy genetyczne. Ich największymi zaletami w stosunku do metod tradycyjnych, jest możliwość pracy z problemami wielowymiarowymi, elastyczność oraz brak konieczności zakładania postaci analitycznej modelu.

### 1.3 Ekonomizacja opieki medycznej

Mimo zauważalnego postępu nauk farmaceutycznych i medycznych we wszystkich ich aspektach, nadal problemem pozostaje koszt i co za tym idzie dostępność nowoczesnych rozwiązań farmakologicznych i procedur medycznych. Zapis w konstytucjach krajów demokratycznych o równości dostępu do zasobów opieki zdrowotnej nie jest realizowany, z powodu kosztów stosowania najnowszych i najlepszych procedur. Jest to spowodowane m.in. przez [96]:

- starzenie się populacji i co za tym spadek odsetka osób w wieku produkcyjnym
- wzrost oczekiwań społeczeństw co do rokowań
- wzrost zachorowań na choroby, których terapia jest kosztowna (nowotwory, AIDS i inne)
- wprowadzanie nowych, kosztowniejszych sposobów leczenia
- wzrost kosztów badań naukowych nad lekami

Wszystkie wymienione powyżej czynniki oraz stały niedobór zasobów finansowych systemów opieki zdrowotnej na świecie, są przyczynami rozwoju stosunkowo nowej dziedziny nauk medycznych jaką jest farmakoekonomika. Przeprowadzenie analizy farmakoekonomicznej umożliwia bezpośrednio porównanie alternatywnych metod terapeutycznych lub diagnostycznych pod kątem ich kosztów oraz efektów.



## 2 Podstawy teoretyczne

### 2.1 Sztuczne sieci neuronowe

#### 2.1.1 Wstęp

Sztuczne sieci neuronowe to jeden z elementów - przeżywającego od końca lat 80-tych drugi okres wielkiej fascynacji - działu nauki, jakim jest inteligencja obliczeniowa. Podstawową różnicą między klasycznymi maszynami liczącymi (komputerami), a układami czerpiącymi z paradygmatu obliczeń neuronowych, jest możliwość przetwarzania równoległego informacji przez układy wielu jednostek podstawowych, wykonujących bardzo prostą instrukcję (sumowanie, przeliczanie przy użyciu zadanej funkcji). Jest to sytuacja skrajnie inna w stosunku do klasycznych procesorów komputerowych opartych na paradygmacie von Neumanna, które najczęściej wykonują kody niezwykle skomplikowanych programów. Równoległość obliczeń i redundancja implikują podstawowe zalety opisywanych systemów w stosunku do klasycznych maszyn liczących:

- odporność na szum w danych
- elastyczność - zdolność do adaptacji
- możliwość samoorganizacji
- możliwość wykonywania obliczeń bez dogłębnej znajomości badanego zagadnienia
- automatyzm pracy sieci neuronowych (łatwość użytkowania)
- możliwość pracy z problemami wielowymiarowymi
- nieliniowy charakter sieci neuronowych umożliwia odwzorowywanie skomplikowanych, nieliniowych zależności

Dzięki wymienionym powyżej właściwościom sztuczne sieci neuronowe stosowane są z powodzeniem w tak różnorodnych i oddalonych od siebie dziedzinach jak:

- chemia [117]
- farmacja [84,118,129]

- botanika [30]
- diagnostyka medyczna [31,36,89]
- stomatologia [26,46]
- ekonomia [3,10,11,143]
- inżynieria i bioinżynieria [9,74]
- geologia [5]

Pierwsze doniesienia na temat koncepcji maszyn potrafiących uczyć się na przedstawianych im danych i generalizować wiedzę pojawiły się w latach 40-tych ubiegłego wieku [55,123], kiedy to McCulloch i Pitts zaproponowali zastosowanie prostego modelu neuronu, realizowanego w postaci elementu progowego. Kolejne ważne etapy w rozwoju teorii sztucznych sieci neuronowych, to praca Hebba badającego mechanizm zapamiętywania informacji na zasadzie adaptacji oraz zbudowanie przez Rosenblatta sieci neuronowej zwanej *Perceptronem*, której jednostki były zorganizowane w warstwy o jednokierunkowych połączeniach od warstwy dolnej do kolejnych [123]. Kolejnym krokiem była sieć Adaline skonstruowana w roku 1960 przez Widrowa i Hoffa.

Zastój obserwowany w dziedzinie prac nad sztucznymi sieciami neuronowymi, który nastąpił po tym odkryciu spowodowany był książką Minsky'ego i Paperta, w której autorzy matematycznie dowodzili niemożności rozwiązania przez wyżej wspomniane sieci jednowarstwowe pewnych nieskomplikowanych obliczeń [88]. Przykładem powyższego jest problem XOR (*exclusive OR*). Mimo, że perceptrony wielowarstwowe mogłyby rozwiązać podobne problemy, to jednak żadna metoda efektywnego uczenia takich struktur nie była znana. Stąd największy wpływ na odrodzenie się tej dziedziny, miało znalezienie algorytmu uczenia sieci, określanego jako propagacja wsteczna (*ang. backpropagation*), dzięki któremu udostępniono sieciom możliwość radzenia sobie z nierozwiązywalnymi do tej pory zadaniami. Autorstwo tej metody przypisywane jest Werbosowi (1974) [137], jednak w piśmiennictwie znaleźć można doniesienia, których autorzy wskazują, że

odkrycia tego dokonało niezależnie kilku badaczy m.in. Bryson i Ho (1969), Parker (1985), Rosenblatt (1985), Hinton (1986) [123].

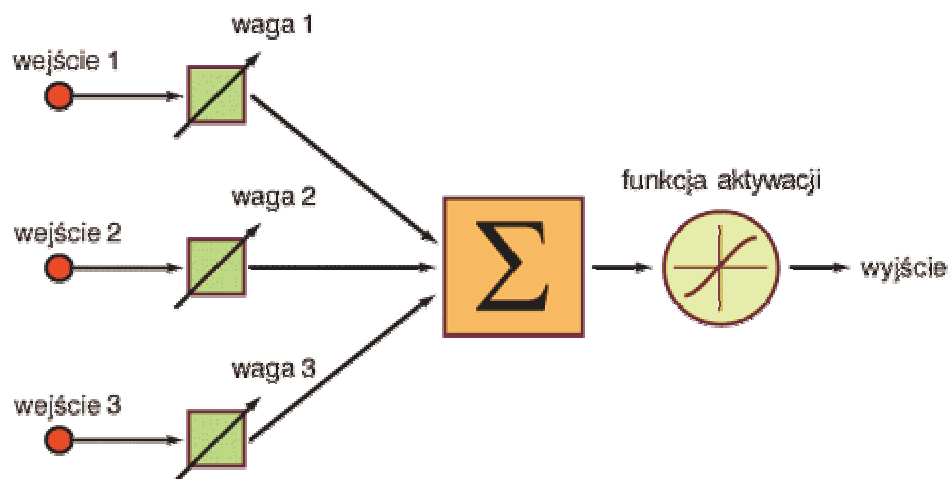
### 2.1.2 Neurobiologiczne inspiracje dla obliczeń neuronowych

Mózg ludzki dzięki swoim właściwościom potrafi dostosować się do zmieniających się warunków zewnętrznych. Problemy określane jako niealgorytmizowalne są przez miliardy ludzi na co dzień rozwiązywane w sposób płynny i naturalny. Złożenie dwóch składowych: szybkości obliczeniowej współczesnych komputerów oraz możliwości płynnego rozwiązywania problemów nieliniowych umożliwiłyby realizację niezwykle ambitnych celów, co stanowi oczywistą zachętę do podjęcia prób maszynowego odwzorowania podobnych zachowań. O ile jednak stan dzisiejszej wiedzy pozwala na w miarę dokładne odwzorowanie działania pojedynczego neuronu, o tyle modelowanie jednoczesnego funkcjonowania kompleksu  $10^{11}$  komórek nerwowych przekracza możliwości dostępnej obecnie technologii. Niemniej jednak sieci neuronowe stanowią jeden z akceptowalnych – choć niezwykle uproszczonych - modeli funkcjonowania ludzkiego mózgu [29,39,72].

### 2.1.3 Podstawowy model sztucznego neuronu i sieci neuronowej

Budowa sztucznego neuronu jest skrajnie uproszczonym modelem neuronu biologicznego.

**Schemat 1** Model sztucznego neuronu wzorowany na neuronie biologicznym.



Podstawowymi elementami tak zdefiniowanego sztucznego neuronu, są [39]:

- wartości docierających do neuronu sygnałów określanych jako wejściowe; mogą to być zarówno informacje pierwotne podawane przez badacza jak i sygnały wyjściowe innej, wcześniejszej komórki w sieci (na Schemacie 1 opisane odpowiednio jako: wejście 1, wejście 2, wejście 3) – odpowiedniki dendrytów w komórce nerwowej
- wagi połączeń międzyneuralnych – kodują siłę, z jaką informacja wejściowa dociera do komórki; wartości wag podlegają modyfikacji w procesie uczenia - w wagach zapisana jest „wiedza” sieci (na Schemacie 1 odpowiednio: waga 1, waga 2, waga 3)
- neuron – odpowiedzialny za obliczanie sumy ważonej wejść (suma ilorazów wartości wejściowych i wartości wag), która to wartość jest następnie przekształcana przez funkcję aktywacji (funkcję przejścia) neuronu; obliczona wartość jest ostatecznym sygnałem wyjściowym neuronu

Dodatkowym elementem nieujęty na powyższym schemacie i w jego opisie jest element progowy (wartość progowa), która jest wielkością graniczną warunkującą powstanie sygnału wyjściowego komórki. W momencie obliczenia sumy ważonej pobudzeń są one porównywane z elementem progowym i tylko w przypadku przekroczenia jego wartości powstaje sygnał wyjściowy dzięki przeliczeniu przez funkcję aktywacji. Pierwotny neuron McCullocha i Pittsa zbudowany był w oparciu o progową funkcję aktywacji (Równanie 1, Schemat 2). Oznacza to, że zasada działania oparta była o regułę *wszystko albo nic* tzn. jeśli wygenerowana wartość pobudzenia była większa od 0 na wyjściu pojawia się wartość 1 i analogicznie, jeżeli wartość pobudzenia nie przekraczała wartości progowej (mniejsza od 0), wtedy na wyjściu pojawiała się wartość 0.

**Równanie 1** Wzór funkcji progowej (jednostkowej).

$$f(x) = \begin{cases} -1, & x \leq x_h \\ 1, & x > x_h \end{cases}$$

gdzie:

$x_h$  – wartość elementu progowego (dla  $x_h=0$  funkcja progowa jest funkcją Heaviside'a)

Ze względu na problemy ujawniające się w trakcie uczenia takich struktur, funkcja progowa jest obecnie rzadko stosowaną funkcją aktywacji [39]. W używanych na co dzień strukturach funkcje aktywacji mają charakter ciągły co oznacza, że wartość wyjścia zmienia się w sposób ciągły (przyjmując również wartości ujemne, w zależności od wartości wag). Przykładem takiej zależności może być prosta funkcja liniowa lecz najczęściej wykorzystywane funkcje mają postać sigmoidy (funkcja logistyczna). Przykładami funkcji logistycznych, które są wykorzystywane w trakcie obliczeń, są unipolarna funkcja sigmoidalna (Równanie 2, Schemat 2) oraz tangens hiperboliczny (Równanie 3).

**Równanie 2** Wzór funkcji sigmoidalnej unipolarnej.

$$f(x) = \frac{1}{1+e^{(-\beta x)}}$$

gdzie:

$\beta$  – zadany parametr

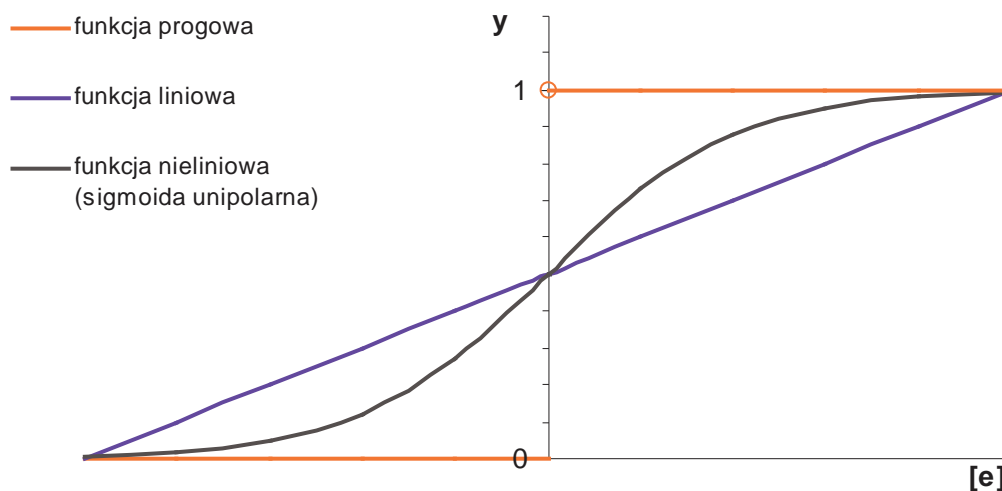
**Równanie 3** Wzór funkcji tangens hiperboliczny.

$$f(x) = \tanh(ax) = \frac{1 - e^{-ax}}{1 + e^{-ax}}$$

gdzie:

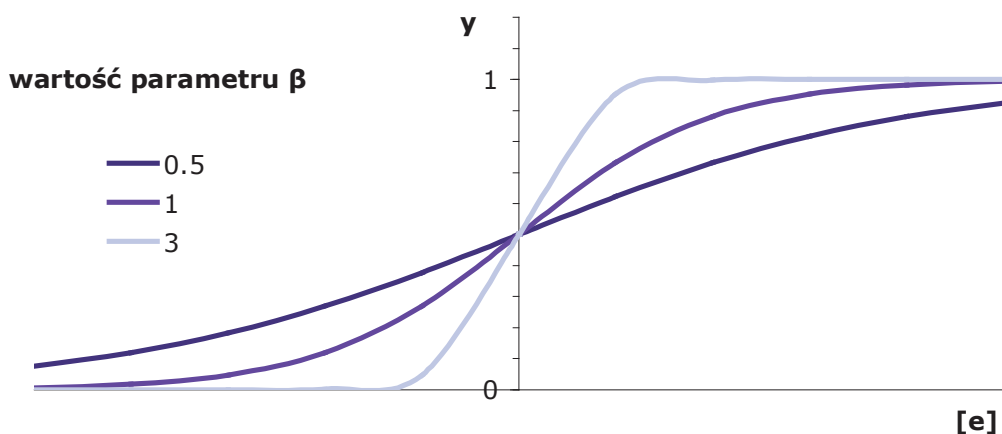
$a$  - zadany parametr

**Schemat 2** Wykresy funkcji progowej (jednostkowej), liniowej i logistycznej (sigmoida unipolarna).



W przypadku funkcji opisywanych przez krzywą logistyczną ich przebieg zależy od parametru (odpowiednio  $\alpha$  oraz  $\beta$  dla sigmoidy oraz tangensa hiperbolicznego).

**Schemat 3** Przebieg funkcji logistycznej w zależności od parametru  $\beta$ .

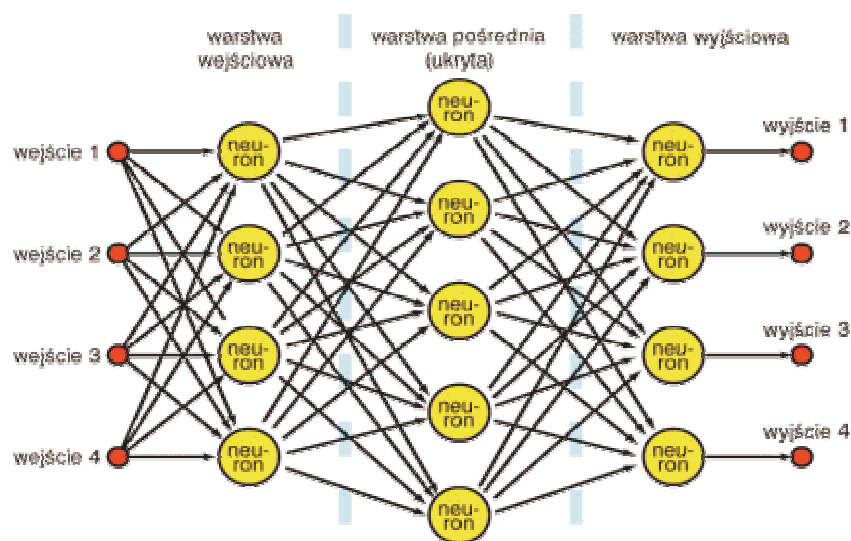


Z powyższych zapisów wynika, że - w przeciwieństwie do układów biologicznych, gdzie istnieje specjalizacja dzieląca komórki i większe struktury na pobudzające i hamujące - jeden neuron może przyjmować zarówno wartości dodatnie (pobudzający) jak i ujemne (hamujący).

Strukturalnie sieci zbudowane są z warstw, z których każda składa się z różnej liczby komórek. Warstwa pierwsza to warstwa wejściowa, która nie wykonuje żadnych obliczeń, a jej zadanie ogranicza się do

podawania sygnału wejściowego. Komórki wejściowe nie są brane pod uwagę przy określaniu liczby warstw sieci i nie są traktowane jako neurony. Stąd najprostsza sieć jest siecią jednowarstwową, gdzie zaraz po komórkach wejścia usytuowane są komórki (lub jedna komórka) wyjściowe. W większości przypadków korzysta się z sieci o strukturze wielowarstwowej, gdzie między dwiema warstwami - wejściową i wyjściową - znajdują się tzw. warstwy ukryte. Zazwyczaj połączenia między neuronami realizowane są na zasadzie każdy z każdym tzn. każda komórka warstwy poprzedniej połączona jest ze wszystkimi komórkami warstwy następnej lecz nie jest to warunek konieczny. Wszystkie komórki muszą mieć dokładnie scharakteryzowaną opisaną powyżej funkcję aktywacji oraz algorytm według którego w procesie uczenia modyfikowane są wagi.

**Schemat 4** Schemat sztucznej sieci neuronowej dwuwarstwowej.

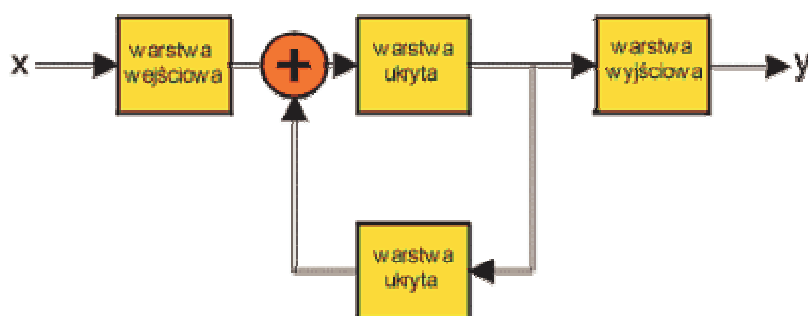


Powyższy schemat (Schemat 4) zakłada jednokierunkowy przebieg sygnałów od komórek wejściowych do wyjściowych. Tego typu sieć określa się niekiedy w piśmiennictwie perceptronami wielowarstwowymi dla podkreślenia jednokierunkowości przebiegu informacji. Cechą charakterystyczną i zaletą tego rodzaju struktur jest stabilne zachowanie w procesie uczenia [39].

Nie jest to jednak jedyna możliwa konstrukcja i oprócz opisanych sieci jednokierunkowych wyróżnić można także sieci, w których informacja

może podlegać sprzężeniom zwrotnym tzn. istnieją połączenia między neuronami późniejszymi i wcześniejszymi (np. sieci rekurencyjne oraz komórkowe) [39]. W sieci neuronowej rekurencyjnej powiązania między komórkami są dowolne, choć najczęściej ma ona budowę warstwową, a przebieg informacji jest dwukierunkowy. W warstwach ukrytych sieci rekurencyjnych zachodzić może sprzężenie zwrotne dodatnie (Schemat 5 – symbol plus) czyli potęgowanie wartości sygnału. Skutkiem krążenia sygnału w sieciach ze sprzężeniem zwrotnym jest możliwość braku stabilności i wystąpienia pewnych skomplikowanych form zachowania (np. w postaci chaosu deterministycznego) [39]. Ten rodzaj sieci jest najczęściej wykorzystywany jako pamięć asocjacyjna np. sieci Hopfielda czy BAM (Bidirectional Associative Memory) [39].

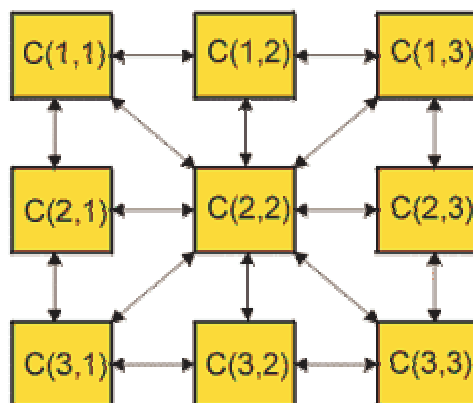
**Schemat 5** Schemat sieci neuronowej rekurencyjnej.



Sieci komórkowe charakteryzują się specyficzną strukturą (Schemat 6). Poszczególne neurony, oznaczone literą C powiązane są z neuronami sąsiednimi (neuron wewnętrzny oznaczony jako C(2,2) najczęściej związany jest z ośmioma najbliższymi sąsiadami), stąd konieczność zdefiniowania pojęcia sąsiedztwa. Tutaj również przebieg informacji jest dwukierunkowy, a podstawowym problemem związanym z tego typu sieciami jest odpowiednie jej zaprojektowanie. Jako przykład tego typu rozwiązania podawana jest sieć typu mapy cech Kohonena [55,123].



**Schemat 6** Schemat sieci neuronowej komórkowej.



#### 2.1.4 Uczenie sieci neuronowych

Jedną z wymienionych wcześniej zalet sieci neuronowych różnicujących je z klasycznymi systemami modelującymi jest umiejętność gromadzenia wiedzy i/lub uczenia się. Działanie nauczonej sieci neuronowej polega na określaniu zmiennej lub zmiennych zależnych na podstawie dostępnych informacji. Powyższa aktywność jest możliwa jedynie w przypadku sieci nauczonych, a więc takich dla których w wyniku procesu uczenia, a więc w sposób adaptacyjny udało się znaleźć zestaw wag połączeń międzyneuronalnych pozwalających odwzorować dane wejściowe na wyjście [98].

Metody uczenia sieci neuronowych dzieli się na trzy podstawowe grupy [98,123]:

- uczenie z nauczycielem (pod nadzorem)
- uczenie z krytykiem
- uczenie bez nadzoru (samoorganizujące się)

Uczenie sieci neuronowej z nauczycielem odbywa się pod nadzorem zewnętrznym. Do sygnałów wejściowych dołączone są pożądane (oczekiwane) wartości wyjścia, a dane podawane są w postaci par – wektor wejściowy/wektor wyjściowy. Przykładem powyższego jest proces minimalizacji funkcji celu [98]. Dla przedstawionej powyżej sytuacji parowania sygnałów wejścia/wyjścia dla każdej pary wyznaczana jest

funkcja błędu (Równanie 4) określana również jako funkcja kosztu lub funkcja celu.

**Równanie 4** Postać funkcji błędu w uczeniu z nauczycielem.

$$e(k) = y(k) - d(k)$$

gdzie:

$y(k)$  – aktualna odpowiedź sieci dla danego wektora wejściowego  $x(k)$

$d(k)$  – zadany wektor wyjściowy dla danego wektora wejściowego  $x(k)$

Bezpośrednim celem uczenia jest minimalizacja funkcji celu (funkcji kosztu –  $e$ ). Najczęściej wykorzystywaną postacią funkcji celu jest błąd kwadratowy (Równanie 5).

**Równanie 5** Postać jednej z funkcji celu – błędu średniokwadratowego.

$$\text{Błąd kwadratowy} = \frac{1}{2} \sum_{i=1}^N (o^p - o^o)^2$$

gdzie:

$i$  – numer wyjścia

$N$  – liczba neuronów wyjściowych

$o^p$  – aktualna wartość wyjścia

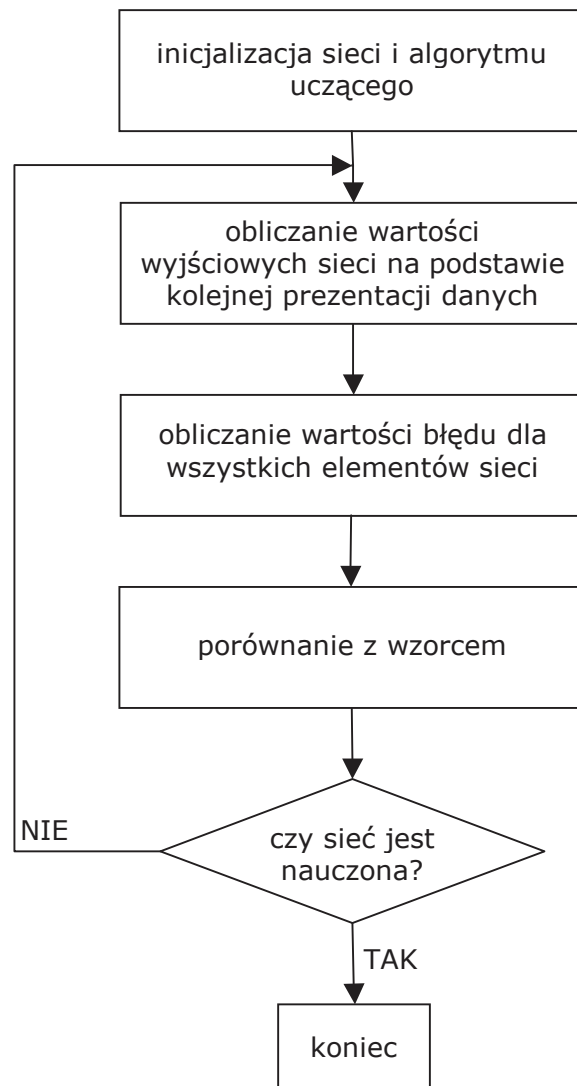
$o^o$  – zadana (wzorcowa) wartość wyjścia

Teoretycznie uczenie prowadzi się aż do momentu kiedy wartość błędu jest równa zeru, jednak w praktyce zadaje się pewną satysfakcjonującą wartość, poniżej której uczenie jest przerywane. Zmiana wag w procesie uczenia może być prowadzona dwojako [39]:

- w sposób przyrostowy – modyfikacja wag następuje po prezentacji każdej pary uczącej, a funkcja błędu zmienia się w każdym kolejnym kroku
- sposób grupowy – zgodnie z zasadą, że błąd całkowity jest sumą błędów dla poszczególnych par wartość wejścia/wartość wyjścia; błąd obliczany jest jako gradient błędu łącznego, a zmiana wagi następuje po prezentacji całego zbioru uczącego

Ogólny schemat przebiegu nadzorowanego uczenia sieci neuronowej przedstawiony jest poniżej (Schemat 7).

**Schemat 7** Schemat blokowy uczenia nadzorowanego sieci.



Uczenie z nauczycielem jest ogólną koncepcją trenowania sieci neuronowej, natomiast podstawowym algorytmem uczenia nadzorowanego sieci wielowarstwowych jest algorytm wstecznej propagacji błędów [39,55,72,123] wraz z modyfikacjami. Proces uczenia polega na prezentowaniu zestawu par wektorów uczący-wzorcowy i odpowiedniej korekcie wektora wag w zadany sposób tak aby zminimalizować popełniany przez system błąd poniżej satysfakcjonującej wartości (metoda gradientowa). Nazwa „propagacja wsteczna” jest pochodną sposobu obliczania wartości błędu – od warstwy ostatniej do

pierwszej, a błąd w warstwie wcześniejszej stanowi funkcję błędów warstwy poprzedzającej [39].

Ogólny, matematyczny schemat uczenia (zmian wartości wektora wag) wielowarstwowej sieci neuronowej przedstawia Równanie 6.

**Równanie 6** Matematyczny zapis zmian wartości wektora wag w zależności od wartości błędu na wyjściu sieci.

$$W(n+1) = W(n) - \eta \nabla E(t)$$

gdzie:

$W$  – wektor wag synaptycznych

$\eta$  – współczynnik szybkości uczenia (najczęściej ustalany *a priori*)

$\nabla E$  - gradient funkcji kosztu

Po przekształceniu w postać skalarną Równanie 6 przyjmuje poniższą postać (Równanie 7).

**Równanie 7** Schemat modyfikacji wag w kolejnych krokach uczenia

$$w_{ij}^{(k)}(n+1) = w_{ij}^{(k)}(n) + \Delta w_{ij}^{(k)}(n)$$

gdzie:

$w_{ij}^{(k)}(n+1)$  - wartość wagi połączenia między i-tym i j-tym neuronem k-tej warstwy w kroku kolejnym

$w_{ij}^{(k)}$  – wartość wagi połączenia między i-tym i j-tym neuronem k-tej warstwy w kroku obecnym

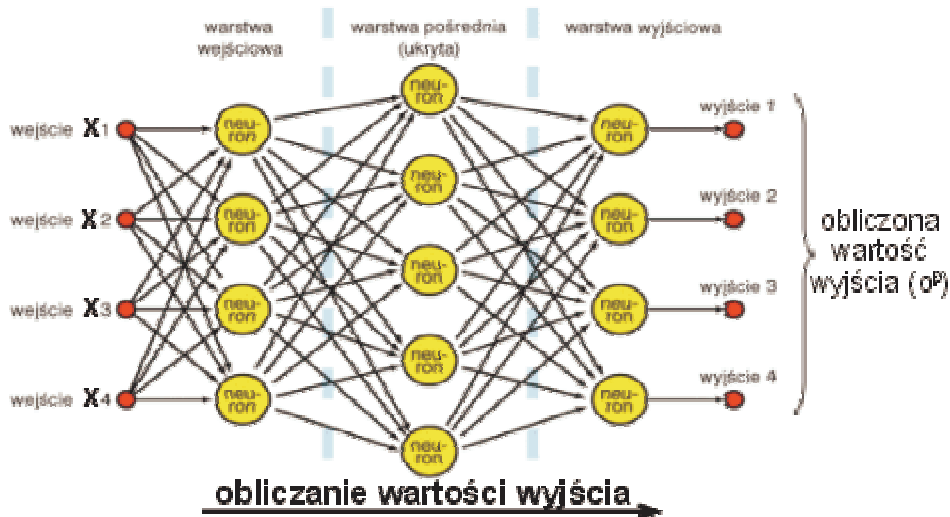
$\Delta w_{ij}^{(k)}$  - zmiana wartości błędu

$n$  – numer kroku

Schematycznie uczenie metodą propagacji wstecznej błędu przedstawiono poniżej (Schemat 8). Kolejny etap to obliczenie wartości błędu ( $b$ ), a więc różnicy między wartością przewidzianą ( $o^p$ ), a oczekiwaną ( $o^o$ ). Następnie dochodzi do wstecznego propagowania wartości błędu przez sieć i odpowiedniej zmiany wartości wag, co przedstawiono na Schemacie 9. Pierwszy etap obejmuje obliczenie

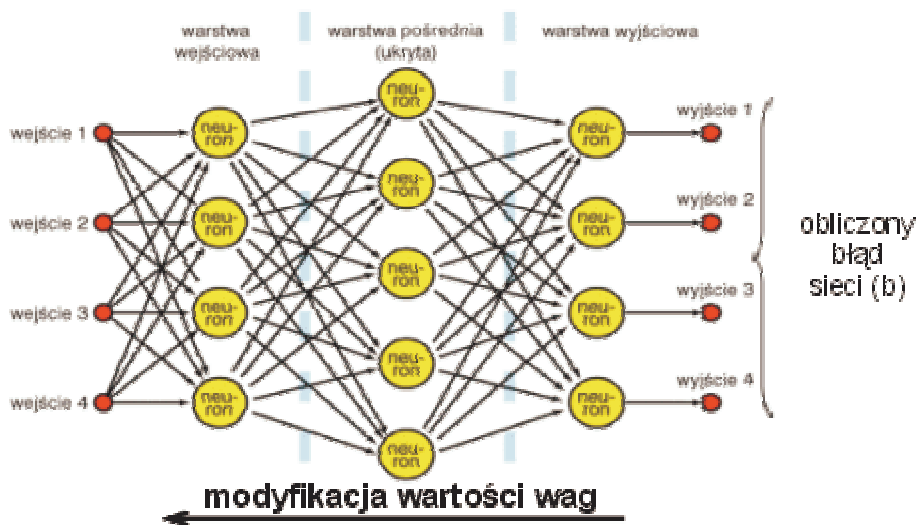
wartości sygnału wyjściowego ( $o^p$ ), po podaniu na wejścia sieci odpowiednich sygnałów wejściowych ( $x_1-x_4$ ).

**Schemat 8** Pierwszy etap uczenia sieci wielowarstwowej metodą wstecznej propagacji błędu.



Kolejny etap to obliczenie wartości błędu ( $b$ ), a więc różnicy między wartością przewidzianą ( $o^p$ ), a oczekiwaną ( $o^o$ ). Następnie dochodzi do wstecznego propagowania wartości błędu przez sieć i odpowiedniej zmiany wartości wag.

**Schemat 9** Drugi etap uczenia sieci wielowarstwowej metodą wstecznej propagacji błędu.



Uczenie sieci ma charakter iteracyjny, a więc opisany schemat powtarzany jest, aż do momentu osiągnięcia zadanego minimum błędu,

ewentualnie – jeżeli minimum nie zostanie osiągnięte – do zadanej maksymalnej ilości iteracji.

Wsteczna propagacja błędu jest systemem uczenia matematycznie niezbyt skomplikowanym, o małej złożoności obliczeniowej oraz umożliwiającym równoległość obliczeń, a przez to uniwersalnym. Pozwala to na jego zastosowanie w szerokim spektrum zagadnień czego pochodną jest duża popularność wśród praktyków wykorzystujących sztuczne sieci neuronowe. Jednocześnie jednak jako algorytm uniwersalny, możliwy do zastosowania zarówno w przypadku problemów błahych jak i niezwykle skomplikowanych, nie jest pozbawiony wynikających z tego wad.

Ceną iteracyjnego charakteru pracy jest czasochłonność procesu uczenia, co mimo możliwości korzystania z coraz wydajniejszych jednostek obliczeniowych pozostaje kłopotliwe.

Główną wadą pozostaje brak pewności osiągnięcia minimum globalnego wielomodalnej funkcji kosztu i możliwość znalezienia przez sieć minimum lokalnego znacznie oddalonego od wartości wzorcowych. Powyższa cecha jest najczęściej wynikiem nieprawidłowego doboru początkowych wartości wag (zbyt duże lub zbyt małe wartości). Oprócz prawidłowego doboru wag, zalecany jest również losowy sposób prezentacji rekordów uczących. Dodatkowo nie ma – jak dotąd – możliwości podania w postaci zapisu matematycznego (funkcyjnego) otrzymanego modelu dlatego modele neuronowe uzyskane w oparciu o algorytm BP określane są mianem „czarnych skrzynek” (*ang. black box*).

Niektóre z wymienionych powyżej niedogodności udało się rozwiązać dzięki licznym modyfikacjom podstawowego algorytmu wstecznej propagacji błędu. Najważniejsze z nich z punktu widzenia praktyki oraz doświadczeń przeprowadzanych na potrzeby niniejszego pracy przedstawiono poniżej.

Współczynnik uczenia  $\eta$  występujący w algorytmie BP może mieć w niektórych sytuacjach dominujący wpływ na spadek gradientu. Jeżeli współczynnik uczenia będzie zbyt mały grozi to zbyt wolnym spadkiem (w przypadku płaskich powierzchni funkcji kosztu) lub w przypadku zbyt

dużego współczynnika  $\eta$  – oscylacjami (fragment powierzchni kosztu o ostrych ścianach) [39,55]. Aby temu zapobiec (przyspieszyć uczenie i zapobiec fluktuacjom) dodano do klasycznego algorytmu tzw. człon momentu, a więc pewną bezwładność (Równanie 8). Zmiana wartości wag zależy także od poprzednich wartości gradientu nadających kierunek, w którym zmierza sieć.

**Równanie 8** Technika momentu.

$$w_{ij}^{(k)}(n+1) = w_{ij}^{(k)}(n) + \Delta w_{ij}^{(k)}(n) + \alpha \Delta w_{ij}^{(k)}(n-1)$$

gdzie:

$w_{ij}^{(k)}(n+1)$  - wartość wagi połączenia między i-tym i j-tym neuronem k-tej warstwy w kroku kolejnym

$w_{ij}^{(k)}$  - wartość wagi połączenia między i-tym i j-tym neuronem k-tej warstwy w kroku obecnym

$w_{ij}^{(k)}(n-1)$  - wartość wagi połączenia między i-tym i j-tym neuronem k-tej warstwy w kroku poprzednim

$\Delta w_{ij}^{(k)}$  - zmiana wartości błędu

$\alpha$  - współczynnik momentu

$n$  - numer kroku

Stała  $\alpha$  dobierana jest z przedziału  $(0,1]$ . Człon momentu pozwala na uniknięcie minimów lokalnych, ponieważ może spowodować chwilowy wzrost wartości funkcji błędu, co pozwala niekiedy na opuszczenie minimum lokalnego [39].

Współczynniki uczenia  $\eta$  i  $\alpha$  mogą mieć zasadniczy wpływ nie tylko na zbieżność procesu uczenia ale również i jego czas. W związku z tym ich prawidłowy dobór pozostaje jednym z najważniejszych problemów przy rozwiązywaniu danego zagadnienia przy użyciu sztucznych sieci neuronowych.

Najprostszy sposób doboru to przyjęcie stałych wartości (z przedziału  $(0,1]$ ), nie podlegających zmianom w trakcie procesu uczenia. Inna metoda zakłada stopniowe zmniejszanie lub wzrost współczynników

w trakcie optymalizacji. Najbardziej zaawansowanym systemem jest reagowanie systemu na wartości błędu sieci i odpowiednia zmiana współczynników  $\eta$  i  $\alpha$  dla każdej z wag. Przykładem powyższej procedury – wykorzystywanej w trakcie doświadczeń – jest reguła delty (*delta-bar-delta*) [39,55,123]. Jej zasada opiera się wprowadzeniu indywidualnego współczynnika uczenia ( $\mu$ ) dla każdej wagi i dodatkowo uzależnienie jego wartości od tendencji zmian gradientu funkcji kosztu. Pozwala to na przyspieszenie procesu uczenia, zwiększenie prawdopodobieństwa osiągnięcia minimum globalnego bez jednoczesnego ryzyka wystąpienia gwałtownych fluktuacji wartości gradientu. Inne algorytmy to m.in. algorytm Quickprop (Fahlmana) oraz RPROP.

Przedstawiona powyżej (Równanie 5) funkcja średniokwadratowa nie jest jedyną możliwą do wykorzystania funkcją błędu. Wymogiem stawianym funkcjom alternatywnym jest różniczkowalność i ciągłość. Wśród funkcji alternatywnych najczęściej wykorzystywane są funkcje Karayannisa, logarytmiczna oraz Hampela.

Zastosowanie funkcji Karayannisa pozwala uniknąć spadkowi tempa zbieżności w ostatniej fazie uczenia z wykorzystaniem funkcji średniokwadratowej [39]. Zaproponowana postać funkcji ma przedstawiony poniżej zapis.

**Równanie 9** Postać funkcji Karayannisa.

$$E(W) = \frac{1}{2} \lambda \sum_{i=1}^N (o^p - o^o)^2 + (1 - \lambda) \sum_{i=1}^N \varphi_1(o^p - o^o)$$

gdzie:

$\lambda$  – współczynnik z zakresu (0,1)

$\varphi_1$  – funkcja błędu (ciągła, dodatnio określona, różniczkowalna)

Stwierdzono, że powyższe kryteria spełnia funkcja, którą przedstawia Równanie 10.



**Równanie 10** Proponowana postać funkcji błędu  $\varphi$  w algorytmie Karayannisa.

$$\varphi_1(x) = \frac{1}{\beta} \ln[\cosh(\beta x)]$$

gdzie:

$\beta$  – dodatnia stała

Istnieją również zależności pozwalające dostosować zmiany wartości współczynnika  $\lambda$  od gradientu błędu.

Funkcja logarytmiczna przyjmująca postać przedstawioną w Równanie 11 pozwala ominąć problem zbieżności procesu uczenia dla danych zaszumianych zawierających rekordy uczące znaczenie odbiegające od średniej (*ang. outliers*) mogących fałszować przebieg funkcji kosztu [98].

**Równanie 11** Alternatywna funkcja celu – funkcja logarytmiczna.

$$E = \sum_{i=1}^N \log\left[1 + \frac{1}{2}(o^p - o^o)^2\right]$$

Podobne zastosowanie ma funkcja Hampela, która dzięki specyficznemu przebiegowi zmian jej pochodnej, stosowana jest w przypadku dużych losowych błędach pomiaru wartości uczących. Funkcję Hampela oraz zawierającą ją funkcję kosztu przedstawiono odpowiednio Równanie 12 oraz Równanie 13.

**Równanie 12** Funkcja Hampela.

$$\phi = \begin{cases} (o^p - o^o)^2 & \text{dla } |o^p - o^o| \leq 1 \\ c_1 \tanh [c_2 (b - |o^p - o^o|)] \operatorname{sgn}(o^p - o^o) & \text{dla } a \leq |o^p - o^o| \leq 1 \\ 0 & \text{dla } |o^p - o^o| > b \end{cases}$$

gdzie:

$a, b$  – zakresy funkcji Hampela

$c_1, c_2$  - stałe

**Równanie 13** Funkcja celu zawierająca funkcję Hampela

$$E = \sum_{i=1}^N \varphi$$

gdzie:

$\varphi$  – funkcja Hampela

Technika *jog of weights* polega na dodawaniu szumu o pewnej niewielkiej wartości do wag sieci, w razie braku poprawy wyników sieci przez zadaną liczbę iteracji (ocena szybkości uczenia sieci) [98]. Liczba iteracji stanowi tzw. „współczynnik cierpliwości” sieci. Dochodzi do zmiany (zaburzenia) wartości wektora wag, co umożliwia sieci opuszczenie minimum lokalnego i przesunięcie do innego punktu funkcji celu. Efektem powyższej aktywności jest niekiedy poprawa skuteczności uczenia sieci i znalezienie lepszego (suboptymalnego) minimum funkcji celu [9,86].

Algorytm wstecznej propagacji błędu wykorzystywany jest w sieciach jednokierunkowych. W przypadku sieci o połączeniach dwukierunkowych lub ze sprzężeniami zwrotnymi konieczne jest zastosowanie innych algorytmów uczenia pod nadzorem [55]. Przykładem powyższych sieci są sieci rekurencyjne. W przypadku sieci o połączeniach symetrycznych, a więc sytuacji gdzie  $\omega_{ij} = \omega_{ji}$  (co jest warunkiem koniecznym) typ sieci i jednocześnie sposób jej uczenia określany jest wspólną nazwą maszyna Boltzmana. Zasadą ich działania jest również poszukiwanie optymalnych wartości wag. Poszukiwanie stanu równowagi odbywa się zgodnie z zasadą rozkładu Boltzmana-Gibbsa (stąd nazwa typu sieci), gdzie podane jest prawdopodobieństwo znalezienia się układu w pewnym określonym stanie [55]. Maszyny Boltzmana stosowane są bardzo rzadko, ze względu na konieczność przeprowadzania żmudnych i długotrwałych obliczeń.

W przypadku metod uczenia ze wzmocnieniem (*ang. reinforcement learning*), w przeciwieństwie do uczenia pod nadzorem, nie są dostępne szczegółowe informacje na temat docelowych wartości wyjścia. Brak ścisłych danych zastąpionych przez wskazówki lub inaczej oceny (np.

informacja binarna – tak/nie) powoduje, że techniki uczenia różnią się. Otrzymywana przez system wskazówka na temat wyjścia nie daje informacji o ewentualnym kierunku (gradiencie) popełnianego błędu. Zakłada się, że sieć uczona ze wzmocnieniem działa w określonym środowisku, które dostarcza informacji na temat podjętych przez sieć działań – wzmocnienia tendencji do określonego zachowania w przyszłości lub jej osłabienie [98].

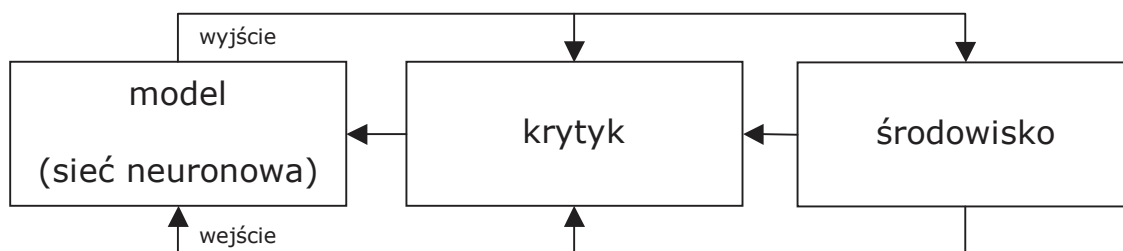
W zależności od natury środowiska uczenie ze wzmocnieniem dzieli się na kilka klas [55]:

- Klasa I – sygnał wzmocnienia jest zawsze taki sam dla par wejście-wyjście
- Klasa II – gdzie występuje tzw. środowisko stochastyczne, a określona para wyznacza jedynie prawdopodobieństwo (o stałej wartości) wzmocnienia
- Klasa III – w której środowisko jest zarządzane dynamicznie, a sygnały wzmocnienia i wartości wejściowe zależą od przeszłych wartości wyjścia (np. teoria gier)

W przypadku problemów klasy I i II najczęściej stosowana procedura uczenia jest algorytm asocjacyjnego nagradzania-karania ( $A_{RP}$ ) [55]. W przypadku problemów klasy III środowisko modelowane jest przez sieć pomocniczą generującą sygnał nagrody/kary (wzmocnienia) dla każdej podanej wartości wyjścia. Schemat 10 przedstawia strukturę funkcjonalną opisywanych systemów. W przypadku klasy III krytykiem może być inna sieć (pomocnicza) modelująca środowisko.

Uczenie z krytykiem oceniane są skutki podjętych akcji i na ich podstawie podejmowane decyzje dotyczące dalszego działania. Dzięki temu jest bardziej uniwersalna niż uczenie pod nadzorem ale jednocześnie realizacja praktyczna jest bardziej skomplikowana [98].

**Schemat 10** Sieć uczona ze wzmocnieniem (z krytykiem).



Główną wadą uczenia z nauczycielem lub z krytykiem jest konieczność dostarczenia systemowi informacji na temat prawidłowej wartości wyjścia lub przynajmniej wskazówek co do odpowiedzi oczekiwanej. Jak podkreślano na wstępie sieci neuronowe są próbą naśladowania działania ludzkiego mózgu w warunkach sztucznych, a oczywistym jest, że ludzki mózg potrafi niekiedy wyciągać wnioski i uogólniać bez dodatkowych informacji na temat natury obserwowanego zjawiska (np. dziecko uczące się rozpoznawać twarze najbliższych) [72]. Wynikające stąd zainteresowanie badaczy konstrukcją systemów uczących się bez dodatkowych informacji zaowocowało algorytmami umożliwiającymi sieciom odkrycie zależności w danych bez żadnych informacji *a priori*.

Uczenie nienadzorowane jest możliwe w przypadku odpowiedniej redundancji (nadmiarowości) danych wejściowych [72]. Zastosowania sieci uczonych bez nadzoru zależą od struktury sieci i stosowanego algorytmu. Klasy rozwiązywanych problemów dzieli się na [55,72]:

- wyszukiwanie podobieństw – wartość wyjścia wygenerowanego przez sieć informuje o stopniu podobieństwa do danej klasy (na ile dany wzorzec jest podobny do innych wzorców w przeszłości)
- analiza czynników głównych – sieć posiada wiele wyjść, z których każdy odpowiada wyróżnionej kategorii; stan aktywności każdego z wyjść odpowiada stopniowi przyporządkowania (na podstawie poprzednio widzianych wektorów)

- klasyfikacja (grupowanie) – wyjścia sieci mają charakter binarny; wyjścia sieci mogą więc przyjąć tylko jeden stan (0 lub 1), a więc pogrupować wzorzec do klas, co jest zadaniem sieci
- poszukiwanie pierwowzoru – sieć również tworzy kategorie ale jednocześnie podaje prototypy danej klasy co umożliwia wykorzystanie jej w postaci pamięci skojarzeniowej (obraz wyjściowy podawany jest jedynie na podstawie wartości wejścia bez dodatkowych informacji z zewnątrz)
- kodowanie – sieć tworzy wektor wyjściowy będący odpowiednikiem wektora wejściowego jednak zapisanego z wykorzystaniem mniejszej ilości bitów; podobne systemy można wykorzystać do transmisji danych przy ograniczonej przepustowości łącza – konieczne jest jednak skonstruowanie odpowiedniego dekodera
- tworzenie map cech – odwzorowanie wzorców wejściowych na geometrycznie uporządkowane wyjście (tablica dwuwymiarowa); idea polega na tworzeniu mapy topograficznej dzięki czemu podobne wartości wejścia powodują pobudzenie bliskich sobie jednostek wyjścia

Co warto podkreślić, powyższe zastosowania nie muszą być traktowane oddzielnie i można tworzyć dowolne ich kombinacje w zależności od rozwiązywanego problemu. Dodatkowo istnieje możliwość tworzenia systemów hybrydowych (połączenia uczenia nadzorowanego z uczeniem bez nauczyciela) co niekiedy może poprawić wyniki lub spowodować redukcję rozmiaru pierwotnego wektora wejściowego (redukcja liczby wymiarów). Efektem tego może być przyspieszenie uczenia z nauczycielem (a właśnie długi czas uczenia jest jedną z jego głównych wad).

Wśród algorytmów uczenia bez nauczyciela wyróżnić można dwie zasadnicze grupy – uczenie konkurencyjne (wykorzystujące uogólnioną regułę Kohonena) oraz uczenie nie oparte o zasadę konkurencji (uczenie asocjacyjne Hebba).

Kolejną grupę stanowią algorytmy uczenia bez nadzoru nie oparte o zasadę konkurencji. Podstawy teoretyczne podobnych algorytmów oparte są o doświadczenia Pawłowa na temat tworzenia się odruchów warunkowych. Opisywane w tym rozdziale metody nie są wykorzystywane do klasyfikacji czy grupowania ale raczej jako narzędzia umożliwiające pomiar podobieństwa. Pierwszą regułą uczenia pojedynczego elementu sieci podał Hebb [72].

Podstawowa reguła podana przez Hebba mówi, że neuron A cyklicznie pobudzany przez inny neuron (B) staje się jeszcze bardziej czuły na jego pobudzenie. Zapis matematyczny powyższej reguły przedstawia Równanie 14 [72].

**Równanie 14** Reguła Hebba.

$$\omega_{AB}(k+1) = \omega_{AB}(k) + \alpha x_A(k)x_B(k)$$

gdzie:

$x_A, x_B$  – stany aktywacji neuronów A i B

$\omega_{AB}$  – waga połączenia synaptycznego między neuronami A i B

$\alpha$  – dodatnia stała

Podstawową wadą prostej reguły Hebba jest brak zbieżności – wagi mogą rosnać bez ograniczeń, a uczenie nigdy się nie kończy. Dodatkowo w opisywanej regule uczonych może być jednocześnie wiele neuronów co powoduje, że jednej klasie wzorców może odpowiadać wiele neuronów, a dla poszczególnych klas zestawy neuronów mogą się pokrywać.

Reguła Oja jest jedną z metod uniknięcia wady pierwotnej reguły Hebba czyli nieograniczonego wzrostu wartości wag. Aby uniknąć złożoności obliczeniowej koniecznej np. przy zastosowaniu prostej normalizacji wag po każdej iteracji, Oja zaproponował modyfikację reguły Hebba, poprzez dodanie modułu zanikania wagi. Dzięki temu zabiegowi algorytm uczenia jest zbieżny.

Dostępne są również inne metody uczenia oparte na przesłankach biologicznych. Wśród technik uczenia sieci bazujących na doświadczeniach

Pawłowa i posiadających źródło biologiczne wyróżnić można m.in. uczenie antyhebbowskie (antykorelacyjne) [98] oraz sieci z modyfikacjami opartymi o doświadczenia Grossberga [72,123].

W algorytmie dekorelacyjnym siła połączenia międzyneuronalnego wzrasta w tym przypadku – odwrotnie w porównaniu z regułą Hebba - gdy sygnały pre- i postsynaptyczne są zdekorowane (odwrotny kierunek pobudzenia). Uczenie antyhebbowskie nie pozwala na nieograniczony wzrost wartości wag, a co za tym idzie nie są konieczne dodatkowe modyfikacje reguły.

Grossberg wprowadził do opisu neuronu dwa kolejne pojęcia – *instar* oraz *outstar*. *Instar* to neuron będący miejscem zbiegania się jednokierunkowych połączeń, natomiast *outstar* jest neuronem wyprowadzającym połączenia neuronalne do innych komórek. Najważniejsze spostrzeżenie dotyczy czasowych zależności aktywności neuronu i wag jego połączeń synaptycznych. Zadaniem neuronu typu instar jest rozpoznawanie wzorców – każdy neuron jest trenowany w celu rozpoznania specyficznego, pojedynczego wektora wejściowego. W efekcie wytrenowany instar jest w stanie rozpoznać zarówno wektor dla którego został wytrenowany jak i wektory do niego podobne (klasa) co jest jednocześnie metodą uniknięcia niedogodności z klasycznej reguły Hebba.

W uczeniu nienadzorowanym typu konkurencyjnego neurony rywalizują aby zostać uaktywnione [98]. Ten typ uczenia określany jest również jako uczenia typu zwycięzca-bierze-wszystko (*ang. winner-takes-all, WTA*), ponieważ pobudzeniu ulega jedynie ten element którego wektor wag jest maksymalnie zbliżony do wektora wejściowego [72]. Oznacza to, że jedynie ten neuron, którego wartość pobudzenia na wyjściu jest największa przyjmuje ostatecznie wartość 1. Co za tym idzie jedynie ten neuron modyfikuje wartości wag do niego dochodzących. Pozostałe neurony przyjmują na wyjściu wartości 0 i nie modyfikują wartości wag. W efekcie współzawodnictwa neurony dopasowują wartości wag w taki sposób aby odpowiadać na jeden typ wektorów wejściowych (klasę wektorów), stąd najczęstszym sposobem wykorzystania tego typu sieci

jest klasyfikacja wektorów. Jednym z najistotniejszych problemów w trakcie uczenia sieci konkurencyjnych jest zdefiniowanie miary odległości między wektorami wejściowymi a wektorem wag neuronu. Najczęściej stosowane to iloczyn skalarny oraz miara euklidesowa.

Powyżej przedstawiono jedynie zarys teorii sieci typu konkurencyjnego. Ich prostota oraz możliwość szerokiego zastosowania spowodowały, że powstało wiele modyfikacji pierwotnej reguły oraz oryginalne typy sieci samoorganizujących się. Wśród nich wymienić można sieci typu ART (*ang. adaptive resonance theory*) bazujące na pracach Grossberga oraz będące hybrydami sieci z kontrapropagacją (sieci Hechst-Nielsen).

Metodą bardzo popularną i dającą w niektórych zastosowaniach bardzo dobre wyniki jest algorytm Kohonena. W klasycznych sieciach Kohonena zaimplementowano opisaną powyżej ideę uczenia jedynie neuronu „zwycięskiego”, a więc takiego, który po prezentacji wzorca uczącego uzyskał maksymalną wartość pobudzenia. Obliczenia rozpoczyna się od małych, przypadkowych wartości wag i po podaniu każdego elementu zbioru wzorców uczących wyznaczany jest neuron zwycięski z zastosowaniem miary euklidesowej. Dodatkowo modyfikowane mogą być również wagi jego otoczenia. Podobna reguła nosi nazwę WTM (*ang. winner-takes-most*), a więc nie oznacza wykluczenia z uczenia wszystkich innych neuronów poza neuronem zwycięskim. Regułę uczenia (algorytm Kohonena) przedstawia Schemat 11.

**Schemat 11** Algorytm Kohonena

$$\omega_i(k+1) = \begin{cases} \omega_i(k) + \alpha(k)(u - \omega(k)) & \text{dla } i \in \Omega \\ 0 & \text{dla } i \notin \Omega \end{cases}$$

gdzie:

$\Omega$  – otoczenie elementu zwycięskiego

$\omega_i$  – wektor wag wstępujących do  $i$ -tego elementu przetwarzającego

$u$  – wektor wejściowy

$\alpha$  – współczynnik uczenia



W sieciach Kohonena najczęściej stosuje się sąsiedztwo gaussowskie, a więc elementy otoczenia neuronu zwycięskiego podlegają zróżnicowanej adaptacji zależnej od wartości funkcji Gaussa.

Algorytm Kohonena nie jest jedynym wykorzystywanym do uczenia tego typu sieci. Inne metody to m.in. algorytm stochastycznej relaksacji, algorytm SCS (*ang. Soft Competition Scheme*) czy algorytm gazu neuronowego [39,55,98].

Testowany w trakcie opisywanych prac algorytm gazu neuronowego polega na zmianie podejścia do pojęcia sąsiedztwa, które tworzy się poprzez uporządkowanie wszystkich neuronów zgodnie z odległością od wzorca. Specyficznie definiowana funkcja sąsiedztwa powoduje, że algorytm gazu neuronowego daje dobre rezultaty szczególnie w przypadku gdy przestrzeń wejść ma skomplikowany nieregularny kształt [39].

Cechą charakterystyczną przedstawionych powyżej metody uczenia sieci neuronowych bez nadzoru jest możliwość zapominania wcześniej nauczonych wzorców w przypadku podawania nowych składników zbioru uczącego. Wymaga to powtarzania procesu uczenia z zastosowaniem nowego, połączanego zbioru [72].

Sieci typu ART (*ang. Adaptive Resonance Theory*) były projektowane jako systemy pozwalające - do pewnych granic - ominąć tę przeszkodę. Teoria rezonansu adaptacyjnego zakłada, że wzorce wejściowe są grupowane do już istniejących klastrów gdy są odpowiednio podobne. W przypadku gdy podawany wzorzec nie jest wystarczająco podobny do żadnego pierwowzoru - tworzony jest nowy z puli „rezerwowych” jednostek wyjściowych [55]. Określenie zakresu podobieństwa do danej klasy dzięki zastosowaniu tzw. parametru czujności (*ang. vigilance*) zawartym w przedziale  $<0;1>$ , który może się zmieniać w trakcie uczenia. Jeśli wartość  $\rho$  jest zbliżona do 1 - powstać może wiele kategorii z niewielkimi różnicami, natomiast niewielka wartość  $\rho$  daje kategoryzację zgrubną [55].

Wyróżnia się trzy typy sieci ART - ART1 dla danych binarnych, ART2 - dla danych ciągłych oraz będąca jej rozwinięciem sieć ART3.

Mimo swoich zalet sieć typu ART ze względu na dużą złożoność obliczeniową (duża liczba koniecznych składników budujących sieć) oraz wrażliwość na zakłócenia [72], jest w praktyce stosunkowo rzadko stosowana.

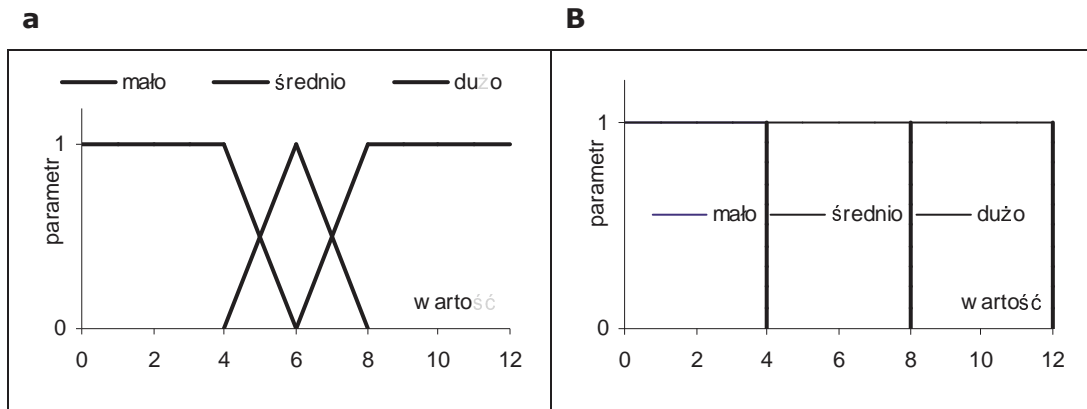
Specyficzną grupą są sieci Hechst-Nielsena tzw. sieci z kontrapropagacją. Sieci typu CP (*ang. CounterPropagation*) są w rzeczywistości sieciami hybrydowymi łączącymi w sobie dwa typy uczenia – bez nadzoru i nadzorowane [39,55,72,123]. Ich cechą charakterystyczną jest obecność – oprócz warstw wejściowej i wyjściowej – również warstwy wewnętrznej działającej na zasadzie konkurencji. Sieci typu CP dzieli się na sieci z pełną kontrapropagacją oraz sieci jednokierunkowe (uproszczone). Warstwa konkurencyjna (tzw. warstwa Grossberga) działa jak klasyczny klasyfikator, przydzielając prezentowany wzorzec do jednej z klas wyznaczonych w pierwszym etapie uczenia zgodnie z zasadą WTA. Drugi etap uczenia to modyfikacja pod nadzorem wag neuronu zwycięskiego łączących go z komórkami warstwy wyjściowej. Sieci typu CP są najczęściej wykorzystywane do szybkiego kojarzenia wektora wejściowego z odpowiednim prototypem [39].

#### 2.1.5 Sieci neuronowe bazujące na logice rozmytej

Dwa z kilku najpoważniejszych zarzutów stawianych klasycznym sieciom neuronowym jako narzędziu wykorzystywanemu do rozwiązywania konkretnych problemów i stosowanych w praktyce, są niemożność analizy postaci modelu (model w postaci niejawnej) oraz brak możliwości operowania informacjami lingwistycznymi. Obydwa z powyższych ograniczeń można ominąć wykorzystując możliwość systemów neuronowych bazujących na logice rozmytej (*ang. fuzzy logic*). Jej twórcą jest prof. Lotfi A. Zadeh, który w swojej pracy [142] przedstawił jej podstawowe założenia. W przeciwieństwie do logiki Boole'a, która klasyfikuje dane w systemie przynależności ostrej, w logice rozmytej ta sama wartość zmiennej wejściowej może zostać zaklasyfikowana do

różnych zbiorów z pewną wartością funkcji przynależności. Graficzną reprezentacją przedstawia Schemat 12.

**Schemat 12** Rozkład funkcji przynależności dla systemu rozmytego (**a**) oraz ostrego (**b**)



Funkcje przynależności stosowane w systemach rozmytych mogą być dowolne jednak najczęściej stosuje się funkcje gaussowskie oraz trójkątne i trapezoidalne [98].

**Równanie 15** Funkcja gaussowska

$$\mu_F(x) = \exp\left(-\left[\frac{x-c}{\delta}\right]^2\right)$$

gdzie:

$\mu_F(x)$  – wartość funkcji gaussowskiej

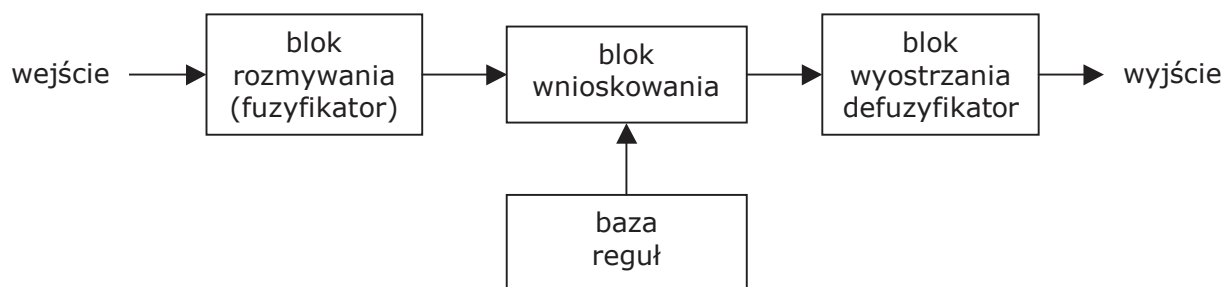
c – centrum funkcji

x – zmienna

$\delta$  – szerokość funkcji

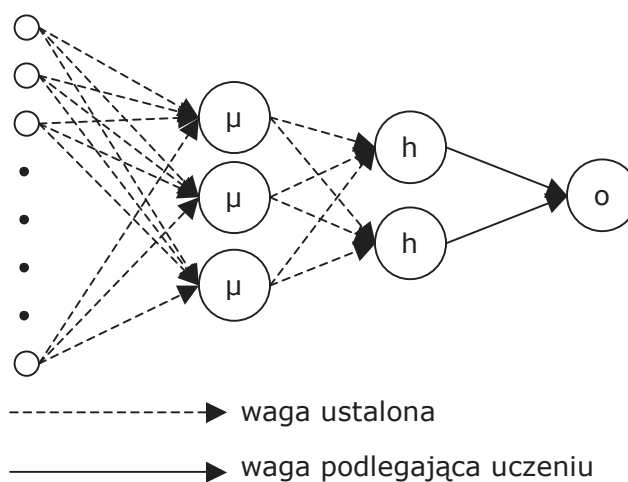
Dane wejściowe to najczęściej informacje przyporządkowane wyjściom w sposób ostry. Aby umożliwić systemowi budowę reguł nieostrych konieczne jest zastosowanie dwóch dodatkowych w stosunku do klasycznych sieci elementów – tzw. fuzyfikatora, przekształcającego dane podawane na wejściu w zbiór nieostry oraz defuzyfikatora, którego zadaniem jest przekształcenie odwrotne (tzn. zebranie powstałych reguł rozmytych i przyporządkowanie im jednoznacznej informacji o charakterze ostrym) [39,98,112]. Przykładowy system rozmyty przedstawiono na Schemacie 13.

**Schemat 13** System rozmyty z blokiem rozmywania (fuzyfikator) oraz wyostrzania (defuzyfikator).



Wyróżnia się dwa zasadnicze typy rozmywania [39]. Najczęściej stosowane jest podejście typu *singleton*, w przeciwieństwie do drugiego – *non-singleton*. Wśród metod defuzyfikacji wymienić można m.in. metodę CA (ang. *Center-Average*), CS (ang. *Center-of-Sums*), COA (ang. *Center-of-Area*), SLIDE (ang. *Semi-Linear Defuzzyfication Method*), M-SLIDE (ang. *Modified Semi-Linear Defuzzyfication Method*), MOM (ang. *Mean-of-Maximum*) [112,141]. Znane są również inne systemy rozmyte np. system Takagi-Sugeno (TSK), który nie wymaga bloku defuzyfikacji [39,141].

**Schemat 14** Graficzne odwzorowanie uczenia sieci neuronowej rozmytej.



gdzie:

$x$  – wejście

$\mu$  – komórka przetwarzająca sygnały zgodnie z funkcją przynależności

$h$  – komórka warstwy ukrytej

W trakcie uczenia sieci rozmytych (sieci o logice rozmytej) można wykorzystać podobne algorytmy w stosunku do klasycznych sieci

neuronowych, w tym wsteczną propagację błędu czy metody samoorganizacji. Schemat 14 przedstawia graficzny zarys uczenia sieci rozmytej z wykorzystaniem metody wstecznej propagacji błędu.

Warstwa pierwsza przetwarza sygnały wejściowe zgodnie z funkcją przynależności  $\mu$ . Wagi łączące komórki opisywanej warstwy z wejściami oraz komórkami warstwy kolejnej nie podlegają procesowi modyfikacji. Uczenie sieci polega na doborze wartości wag łączących komórki kolejnej warstwy ukrytej z warstwą wyjściową ( $\omega$ ) oraz parametrów  $c$  i  $\delta$ .

#### 2.1.6 Sieci neuronowe wykorzystujące algorytmy genetyczne

Zastosowanie w analizie neuronowej algorytmów bazujących na regułach znanych z genetyki, ma trzy główne kierunki [138]. Pierwszym z nich jest modyfikacja wag czyli wykorzystanie do uczenia sieci. Kolejne to kształtowanie topologii (architektury sieci) oraz preprocessing danych uczących (m.in. redukcja rozmiaru wektora wejściowego) i analiza zachowań sieci (manifestowanych w postaci zmian wartości wyjścia). Podstawą algorytmów genetycznych są zjawiska znane z genetyki, których celem jest dostosowanie organizmu (w tym przypadku – sieci neuronowej) do zmieniającego się środowiska. Szczegółowy opis reguł genetycznych wykracza poza ramy tej pracy.

#### 2.1.7 Przygotowanie danych uczących i testowych oraz dobór architektury sieci i współczynników uczenia

Wykorzystując sieci neuronowe jako narzędzia do rozwiązywania pewnych problemów po etapie jego definiowania następuje etap przygotowania danych uczących oraz projektowania struktury sieci[98].

Dane wejściowe wykorzystywane jako informacje uczące dla sieci mają zasadnicze znaczenie dla jakości uzyskiwanego wyniku. Podstawowym warunkiem jaki muszą spełniać jest reprezentatywność dla danego problemu. Sposób przetwarzania danych dla sieci będzie w głównej mierze zależał od charakteru rozwiązywanego problemu, a co za tym idzie od koniecznego sposobu prezentacji informacji uczącej. Wśród

możliwych typów przekształceń wymienić można m.in. następujące operacje [39]:

- filtracja – usunięcie niezamierzonego szumu i zakłóceń z danych pierwotnych
- zaszumianie danych – dodanie losowego szumu do danych pierwotnych może w pewnych przypadkach pozytywnie wpłynąć na wyniki generalizacji
- binaryzacja – zmiana wartości wejścia o charakterze ciągłym w wartości binarne (najczęściej za pomocą progowania - zasadnicze znaczenie w przypadku zastosowania tej techniki ma dobór wartości progu)
- skalowanie – dopasowanie wymiarów wektora pierwotnego
- normalizacja – rozciągnięcie zakresu wartości atrybutu do szerokości założonego przedziału (skalowanie i normalizacja może odbywać się względem wartości maksymalnej, średniej, odchylenia od średniej, odchylenia od wartości minimalnej czy względem standardowego odchylenia średniokwadratowego)
- analiza składowych głównych – PCA (*ang. Principal Component Analysis*)
- balansowanie zbiorów – wycinanie lub multiplikowanie losowych rekordów zbioru uczącego aż do momentu wyrównania liczebności klas

Wymienione powyżej techniki nie opisują wszystkich możliwych, a jedynie najczęściej spotykane lub wykorzystane w trakcie doświadczeń.

Nauczona sieć neuronowa powinna cechować się zdolnością do generalizacji wiedzy, a więc jej uogólnianiem. Aby stwierdzić powyższą zdolność, konieczny jest podział zgromadzonych informacji na zbiory uczące – wykorzystywane w trakcie uczenia modelu oraz testowe – wykorzystywane do oceny skuteczności jej działania. Dostępne techniki obejmują podział na dwa (uczący-testowy) lub trzy (uczący-walidujący-testowy) podzbiory. Najczęściej wykorzystywana metoda podziału to tzw. wzajemne sprawdzanie (*ang. cross-validation*) [40,83]. Wśród dostępnych

możliwości w literaturze dominuje 10-cio krotne wzajemne sprawdzanie (*ang. 10-fold cross-validation*) lub technika *leave-one-out*.

Mimo istniejących algorytmów budowy architektury sieci jak i modyfikacji w trakcie uczenia, a nawet technik automatycznego tworzenia [55], najskuteczniejszą metodą poszukiwania zarówno samej architektury jak i doboru parametrów uczenia pozostaje metoda prób i błędów.

#### 2.1.8 Analiza nauczonego modelu neuronowego

Mimo, że sieci neuronowe określane są mianem czarnej skrzynki, uniemożliwiającej podanie postaci analitycznej modelu istnieją możliwości analizy budowy sieci i jej redukcji. Jedną z nich jest analiza wpływu poszczególnych wartości podawanych na wejściu na uzyskiwany wynik. Czynniki składowe wektora wejściowego mogą zawierać informacje, które w minimalnym stopniu biorą udział w tworzeniu sygnału wyjściowego. Niewielka dawka informacji niesiona przez poszczególne jego składowe może powodować, że wskazane staje się ich pomijanie [98]. Wytłumaczenie powyższego zjawiska wiąże się z określeniem opłacalności obliczeniowej. Modelowanie złożonych zjawisk, kiedy rozmiar wektora wejściowego jest bardzo duży, jego redukcja może spowodować drastyczne zmniejszenie zapotrzebowania na moc obliczeniową. Przy relatywnie niewielkim wzroście wartości błędu (niekiedy nawet jego redukcji) okazuje się to opłacalne.

Stosowana w trakcie doświadczeń metoda redukcji ilości wejść oparta została na analizie wrażliwości sieci względem danych uczących [98], przeprowadzanej po wstępnym etapie uczenia. Po określeniu poszczególnych składników wrażliwości wyjścia względem wejścia określane są tzw. współczynniki ważności. Eliminacji części wejść można dokonać, jeśli zaznacza się wyraźna różnica we wspomnianych powyżej współczynnikach między dwoma kolejnymi składnikami w szeregu (składniki poniżej). W przypadku gdy rozkład spadku ich wartości jest równomierny, nie ma podstaw do wyciągania wniosków na temat ich mniejszego wpływu na wartość wyjściową.

Oprócz przedstawionej powyżej techniki istnieje wiele innych, wśród których wymienić można m.in. metodę NSA (Numeric Sensitivity Analysis) [90] czy też technikę zaproponowaną przez Hashema [54].

## 2.2 Farmakoekonomika

### 2.2.1 Wstęp

Pierwotne podejście do problemu oceny leków ograniczało się jedynie do oceny ich skuteczności. Preparaty stosowane w leczeniu niektórych schorzeń, które w swoich czasach stanowiły przełom terapeutyczny obecnie nie zostałyby dopuszczone nawet do etapu badań klinicznych. Przykładem podobnych środków może być chociażby słynny „preparat 606” czyli Salwarsan – chemioterapeutyk stosowany w leczeniu kiły zsyntetyzowany w 1909 roku przez Paula Ehrlicha. Działania niepożądane leku wpływające na bezpieczeństwo farmakoterapii, nie były badane przez długi czas stosowania chemicznych środków leczniczych. Dopiero doświadczenia ostatnich dziesięcioleci i tragedie wywołane przez niektóre leki, zwróciły uwagę na drugi aspekt terapii lekami, jakim jest jej bezpieczeństwo.

Pełna analiza farmakoekonomiczna jest trzecim, uzupełniającym ogniwem oceny prowadzonej terapii. Celem analizy farmakoekonomicznej jest ocena zarówno efektów jak i kosztów leczenia oraz porównanie stosowanych procedur. Jako, że pełna analiza farmakoekonomiczna jest analizą porównawczą, przy jej wykonywaniu mogą być brane pod uwagę następujące alternatywy:

- porównanie dwóch lub więcej metod farmakoterapii
- porównanie farmakoterapii i innych sposobów leczenia (np. leczenie zabiegowe)
- porównanie wyników i kosztów leczenia z analogicznymi wskaźnikami przy zaniechaniu leczenia

Nie ma możliwości przeprowadzenia analizy farmakoekonomicznej dla wszystkich nowowprowadzanych schematów terapeutycznych, ani tym bardziej dla terapii stosowanych od dawna. Analiza farmakoekonomiczna



powinna być przeprowadzana wtedy, gdy mimo wysokich kosztów generowanych w trakcie wdrażania programu można spodziewać się znaczących efektów.

Wykonanie analizy farmakoekonomicznej pozwala oprzeć się na racjonalnych przesłankach dotyczących opłacalności wdrażania danej metody leczenia w stosunku do porównywanych alternatyw. Celem jej wykonywania jest dostarczenie gremiom podejmującym decyzje o charakterze ekonomicznym w ochronie zdrowia[23]. Przykładem obrazującym powyższe założenie mogą być wymagania przeprowadzenia analizy o charakterze ekonomicznym w trakcie procesu rejestracji leku lub wpisania na listę leków refundowanych w niektórych krajach.

Podkreślenia wymaga fakt, że farmakoekonomika nie jest jedynie narzędziem umożliwiającym oszczędzanie, kosztem jakości stosowanej terapii. Podstawowym celem jej stosowania jest racjonalizacja wydatków i alokacja zasobów na podstawie wykonanych analiz. Nie jest to jednak ostateczny wskaźnik, jako że pod uwagę musi być brany również czynnik etyczny, przykładem którego są choćby tzw. leki sieroce (*ang. orphan drugs*), jednak podobne rozważania nie wchodzą w zakres niniejszej pracy.

### 2.2.2 Etapy badania farmakoekonomicznego

Zgodnie ze światowymi wytycznymi [23,94,96] prawidłowo przeprowadzona analiza farmakoekonomiczna powinna być zbudowana z następujących etapów:

- zdefiniowanie problemu, określenie odbiorców wyników badania
- określenie i zdefiniowanie alternatywnych metod terapeutycznych, łącznie z dokładnym zdefiniowaniem docelowych grup pacjentów dla analizowanych leków
- wyznaczenie perspektywy analizy
- określenie ram czasowych badania oraz określenie horyzontów czasowych oceny konsekwencji umożliwiających objęcie całości skutków stosowania ocenianej procedury medycznej

- dobór odpowiedniego typu analizy (techniki analitycznej)
- identyfikacja i ocena kosztów i konsekwencji na podstawie wiarygodnych danych
- dyskontowanie kosztów i konsekwencji
- analiza inkrementalna
- przedstawienie wyników
- analiza wrażliwości

### 2.2.3 Źródła danych wykorzystywanych w analizach farmakoekonomicznych

Każda ocena ekonomiczna i wyciągane na podstawie jej wyników wnioski – bez względu na przedmiot oceny – są na tyle wartościowe, na ile reprezentatywne i odpowiednie były dane wykorzystane w trakcie jej przeprowadzania. Jednocześnie jednak analitycy zmuszeni są korzystać z takich informacji, jakie są w danej chwili dostępne, przeprowadzać ich wnikliwą analizę oraz na ich podstawie wyciągać możliwie najdokładniejsze wnioski [13,38].

Informacje konieczne do przeprowadzenia odpowiedniej analizy podzielić można na dwie grupy:

- dotyczące środków farmaceutycznych lub technologii medycznych podlegających analizie, w tym kosztów związanych z ich stosowaniem
- dotyczące efektów klinicznych związanych z ich stosowaniem (wyników)

Najczęściej wykorzystywane dwie metody zbierania danych wykorzystywanych w trakcie analiz farmakoekonomicznych obejmują obserwacje równoległe z badaniami klinicznymi (ang. piggy-back study) oraz prowadzone specjalnie pod kątem wykonania analizy farmakoekonomicznej. Podstawowa różnica dotyczy tzw. wiarygodności wewnętrznej czyli możliwości dostarczenia niezafałszowanej informacji na temat skuteczności klinicznej i co za tym idzie wyciągnięcia wniosków na temat zależności przyczynowo-skutkowych oraz zewnętrznej czyli

możliwości ekstrapolowania wyników na populację szerszą niż populacja badana. Informacje uzyskiwane z randomizowanych, kontrolowanych, zaślepianych badań klinicznych cechują się wysoką wiarygodnością wewnętrzną, natomiast istnieje ryzyko, że ich wyników nie będzie można rozciągnąć na całą populację. Badania obserwacyjne z kolei cechują się wysoką wiarygodnością zewnętrzną, natomiast wyciągane na ich podstawie zależności przyczynowo-skutkowe mogą okazać się fałszywe. Proponowana w polskich wytycznych przeprowadzania badań farmakoekonomicznych hierarchia źródeł jest następująca [103]:

- a) badania eksperymentalne
  - randomizowane, kontrolowane badania kliniczne
  - kontrolowane badania kliniczne z pseudo-randomizacją
  - kontrolowane badania kliniczne bez randomizacji
- a) badania obserwacyjne
  - badania kohortowe prospektywne z równoległą kontrolą
  - badania kohortowe prospektywne z historyczną kontrolą
  - badania kohortowe retrospektywne z równoległą kontrolą
  - badania epidemiologiczne kliniczno-kontrolne retrospektywne
  - badania typu „przed-po”
- a) opinie ekspertów
  - opinie uznanych ekspertów (raporty komitetów ekspertów, badania opisowe)

Mimo trudności łączenia informacji z różnych źródeł podkreśla się konieczność dopełniania danych z badań klinicznych badaniami obserwacyjnymi [38]. W przypadku zastosowania techniki metaanalizy konieczne jest zastosowanie przyjętych wytycznych dotyczących jej wykonywania oraz podanie kryteriów doboru włączonych badań.

Opinie ekspertów nie zastępują dowodów naukowych i powinny być wykorzystywane jedynie w wypadku braku wyników badań w celu określenia kontekstu dla oceny farmakoekonomicznej [103].

#### 2.2.4 Typy analiz farmakoekonomicznych

W literaturze fachowej wymienianych jest pięć podstawowych typów analiz farmakoekonomicznych:

- analiza minimalizacji kosztów (ang. cost-minimisation analysis, CMA)
- analiza efektywności kosztów określana według dawnej nomenklatury jako analiza koszt-efektywność (ang. cost-effectiveness analysis, CEA)
- analiza użyteczności kosztów określana według dawnej nomenklatury jako analiza koszt-użyteczność (ang. cost-utility analysis, CUA)
- analiza wydajności kosztów określana według dawnej nomenklatury jako analiza koszty-korzyści (ang. cost-benefit analysis, CBA)
- analiza koszt-konsekwencje (ang. cost-consequence analysis, CCA)

Dodatkowo wymieniana jest również analiza kosztów [16,96], która nie jest pełną analizą farmakoekonomiczną, jako że dotyczy jedynie fragmentu wybranego z całości koniecznych do zanalizowania danych. Podczas opracowywania informacji nie są brane pod uwagę wyniki stosowanych terapii. Schematyczny opis poszczególnych typów analiz zawiera Tabela 1.

**Tabela 1** Charakterystyka typów analiz farmakoekonomicznych oraz analizy kosztów.

<b>Typ analizy</b>	<b>Jednostka pomiaru kosztów</b>	<b>Porównywane wyniki terapii</b>	<b>Jednostki pomiaru wyników terapii</b>
Analiza kosztów	Jednostki pieniężne	-	-
Analiza minimalizacji kosztów	Jednostki pieniężne	Muszą być identyczne dla porównywanych programów terapeutycznych	-
Analiza efektywności kosztów	Jednostki pieniężne	Muszą być mierzone z wykorzystaniem tej samej jednostki, co umożliwia bezpośrednie porównanie	Naturalne (wyniki kliniczne) np. mmHg ciśnienia tętniczego, zyskany okres życia (tygodnie, lata).
Analiza użyteczności kosztów	Jednostki pieniężne	Jakość życia chorego	QALY ( <i>ang. quality adjusted life year</i> ) - skorygowane o jakość lata życia pacjenta HYE ( <i>ang. healthy-year equivalent</i> ) – odpowiednik roku życia w pełnym zdrowiu SAVE ( <i>ang. saved-young-life equivalent</i> ) – równoważnik uratowanej od śmierci jednej młodej zdrowej osoby
Analiza wydajności kosztów	Jednostki pieniężne	Mierzone różnymi jednostkami naturalnymi, które następnie przeliczane są na jednostki monetarne	Jednostki monetarne
Analiza koszt-konsekwencje	Jednostki pieniężne	Całość obserwowanych wyników terapii (zarówno kliniczne jak i jakość życia)	Jednostki kliniczne (naturalne) oraz jednostki jakości życia (QALY, HYE, SAVE)

#### 2.2.4.1 Analiza kosztów

Analiza kosztów (*ang. cost analysis*) oraz analiza kosztów choroby (*ang. cost-of-illness analysis*) nie jest pełną analizą farmakoekonomiczną, ponieważ nie są brane pod uwagę konsekwencje stosowanych procedur. Narzędzia pomiarowe, na których bazuje analiza kosztów wykorzystywane są analogicznie we wszystkich innych typach analiz farmakoekonomicznych [38].

Koszt definiowany jest jako wymiar nakładów poniesionych w trakcie stosowania danego programu zdrowotnego [38,96], w celu uzyskania określonego wyniku. Koszt nie może być utożsamiany z ceną leku lub danej procedury medycznej, wydatkami pacjenta lub świadczeniodawcy, a jego ocena jest złożona. Koszty w opiece medycznej podzielić można wielorako:

##### a) koszty bezpośrednie, pośrednie, niewymierne

Koszty bezpośrednie są w jednoznaczny sposób związane z prowadzonym programem terapeutycznym. Dzielone są na koszty medyczne (leki, aparatura medyczna, praca personelu specjalistycznego, hospitalizacja, procedury diagnostyczne) oraz niemedyczne (koszty transportu do szpitala lub ambulatorium, specjalistyczna dieta, koszty opieki nad chorym).

Koszty pośrednie nie są bezpośrednio związane z chorobą ale stanowią jej finansowe konsekwencje (utrata dochodów osoby chorej związana z nieobecnością w pracy lub obniżeniem produktywności, utrata sprawności fizycznej, umieralność).

Koszty niewymierne są najtrudniejszymi w ocenie, jako że oparte są na subiektywnej ocenie niemożliwych do bezpośredniego przeliczenia na wartości pieniężne wrażeń chorego (ból, cierpienie, spadek jakości życia).

##### b) koszty stałe i zmienne

Koszty stałe nie zmieniają się bez względu na wynik kliniczny osiągnięty w danym programie (np. amortyzacja szpitala, sprzętu), w

przeciwieństwie do drugiego rodzaju kosztów, których wielkość zależy od ilości chorych i rezultatu terapii (np. koszty farmakoterapii).

c) inne rodzaje kosztów w farmakoekonomice

Inne rodzaje kosztów wymieniane w analizie ekonomicznej to m.in. koszt średni, inkrementalny i marginalny. Koszt średni definiowany jest jako wielkość nakładów finansowych zużytych na uzyskanie jednostki efektu medycznego. Koszt marginalny jest zmianą kosztu całkowitego powstałą w wyniku uzyskania dodatkowej jednostki efektu, natomiast koszt inkrementalny to różnica w kosztach uzyskania jednostki efektu w przypadku dwóch lub więcej procedur podlegających porównaniu.

Podobnie jak w pełnej analizie farmakoekonomicznej, tak i w analizie kosztów konieczne jest zdefiniowanie perspektywy, z jakiej dokonywana jest ocena, jako że inne znaczenie i wartości będą miały koszty oszacowane z pozycji pacjenta, świadczeniodawcy, płatnika czy społeczeństwa. Co za tym idzie wnioski wyciągane na ich podstawie mogą być różne. Ocena kosztów obejmuje szacowanie całości zasobów wykorzystanych w czasie przeprowadzania opisywanego programu zdrowotnego (terapeutycznego, profilaktycznego), w odpowiednich jednostkach dla poszczególnych składowych. Ich cechą wspólną jest możliwość przeliczenia uzyskanych wartości na jednostki pieniężne. Perspektywą najszerszą i polecaną jako właściwa jest perspektywa społeczna [38].

W trakcie przeprowadzania analizy kosztów jednym z zasadniczych czynników mogących wpływać na wyniki końcowe jest okres monitorowania kosztów [38]. Zbyt krótki okres poddany analizie może skutkować niedoszacowaniem kosztów jednej lub większej ilości zastosowanych procedur, co z kolei może wypaczyć końcowe wyniki pełnej analizy farmakoekonomicznej.

Etapami kończącymi analizę kosztów powinny być dyskontowanie oraz analiza wrażliwości. Potrzeba dyskontowania oparta jest na założeniu spadku wartości pieniądza w czasie i jednocześnie konieczności oceny odległych w czasie skutków poszczególnych programów. Co za tym idzie

proces dyskontowania ma umożliwić porównanie potencjalnych kosztów ponoszonych w przyszłości, z kosztami teraźniejszymi dzięki pomniejszeniu ich wartości o tzw. współczynnik dyskontowania, którego wartość zależy od wysokości stopy dyskontowej oraz okresu czasu, dla którego liczone są odpowiednie wartości. Analiza wrażliwości ma za zadanie ocenić wpływ poszczególnych składowych kosztów na wynik końcowy, przy użyciu prostej techniki drastycznej zmiany ich wartości (próg 50%). Jako zmienne wrażliwe definiowane są te elementy składowe całości, których modyfikacja w sposób drastyczny wpłynie na końcową wartość kosztu.

#### 2.2.4.2 Analiza minimalizacji kosztów

Przeprowadzanie analizy minimalizacji kosztów jest zasadne tylko w sytuacji, gdy wynik zdrowotny porównywanych schematów jest identyczny. Porównanie dwóch lub więcej programów ogranicza się do analizy kosztów, a celem jest znalezienie najtańszej alternatywy. Podstawową trudnością i ograniczeniem przedstawianego typu analizy jest konieczność przyjęcia braku różnic terapeutycznych, co jest możliwe jedynie na podstawie uzasadnionych danych [96,111]. Analizę kosztów przeprowadza się w sposób analogiczny do przedstawionego w rozdziale 0.

Preferowaną alternatywą jest ta, której przeprowadzenie związane jest z mniejszymi kosztami całkowitymi. Przykładem stosowania analizy minimalizacji kosztów są m.in. badania dotyczące oceny kosztowej schematów diagnostycznych, których zakładana skuteczność jest identyczna [114], oceny kosztów opieki nad chorym w warunkach szpitalnych oraz w domu [70,119], a także kosztów generowanych w trakcie farmakoterapii chorób nowotworowych różnymi schematami terapeutycznymi, których skuteczność jest równa, czego przykładem może być analiza terapii niedrobnokomórkowego raka płuc przedstawiona w pracy [11].



#### 2.2.4.3 Analiza efektywności kosztów (koszt-efektywność)

Analiza efektywności kosztów to przykład pełnej analizy farmakoekonomicznej, gdzie szacowane są zarówno koszty jak i efekty porównywanych schematów terapeutycznych. Możliwe jest również porównanie kosztów i efektów stosowanego schematu terapeutycznego z sytuacją zaniechania leczenia. Podstawowym warunkiem umożliwiającym wybór analizy koszt-efektywność do oceny interesujących badacza programów jest zastosowanie tych samych jednostek pomiaru wyników [48,96]. W przeciwieństwie do pozostałych typów analiz wykorzystywane w tym typie analizy jednostki zaliczane są do wartości klinicznych - zmiana ciśnienia tętniczego krwi - mmHg, zmiana stężenia glukozy we krwi - mmol/l lub mg% lub naturalnych - wyleczenie, zyskany okres życia np. zyskane lata życia (*ang. life years gained*), ilość dni wolnych od objawów. Polskie wytyczne przeprowadzania badań farmakoekonomicznych zalecają przetransponowanie wyników pierwotnych, wyrażonych w wartościach klinicznych (tzw. surogaty) na jednostki naturalne [103]. Stosowane są również miary specyficzne dla ocenianych leków np. częstość ataków astmy w przypadku preparatów przeciwastmatycznych. Dzięki swojej elastyczności opisywany typ analizy może być wykorzystywany nie tylko w przypadku oceny schematów lekowych ale w każdym przypadku kiedy wyniki porównywanych opcji są wyrażane w tej samej jednostce. Stąd możliwość zastosowania w przypadku oceny programów o charakterze profilaktycznym [110] lub badań przesiewowych [52].

Analiza efektywności kosztów powinna być przeprowadzana zgodnie ze schematem przedstawionym w rozdziale 2.2.2. Najczęściej wykorzystywaną perspektywą badania jest perspektywa społeczna, niemniej jednak możliwe jest także przeprowadzenie równoległych analiz z punktu widzenia innego odbiorcy wyników. Określenie jednostek efektywności wykorzystanych dla ocenianych schematów (różne w zależności od problemu) musi w najlepszy możliwy sposób odzwierciedlać zmiany zdrowotne obserwowane w trakcie ich stosowania oraz umożliwić

zaobserwowanie ewentualnych różnic między programami. Projekt polskich wytycznych przeprowadzania analiz farmakoekonomicznych [103] nakłada na prowadzącego badanie, obowiązek zdefiniowania i opisanie wszystkich alternatyw terapeutycznych oraz określenia alternatywy (lub alternatyw) referencyjnej. Najczęściej jest to schemat leczenia, który może być zastąpiony przez badany lek czyli leczenie pierwszego rzutu. Inne dopuszczalne alternatywy to praktyka najczęściej stosowana, o największej skuteczności oraz o najniższym koszcie. Dopuszcza się również porównanie z placebo (nieleczenie).

Analiza kosztów powinna być przeprowadzana zgodnie ze schematem podanym w rozdziale 0. Analiza efektywności kosztów wymaga przedstawienia efektu w postaci liczbowej. Efektywność jest określana jako zdolność do osiągnięcia poprawy stanu zdrowia w realiach praktycznych [48]. Wynika stąd, że efektywność rozumiana jest inaczej niż skuteczność (efekt uzyskiwany w sztucznie stworzonych warunkach idealnych) oraz stosowność, która jest pojęciem szerszym i bierze pod uwagę skutki niepożądane, praktyczne możliwości zastosowania oraz koszty. Efektywność stanowi wypadkową obu przedstawionych zmiennych. Obecność licznych czynników wpływających na efekt utrudnia jego pomiar. Nawet dla najlepiej zaplanowanych badań i doświadczeń niekiedy trudno jest określić wpływ poszczególnych z nich na uzyskiwany wynik. Jednym z najważniejszych czynników jest - opisywany powyżej - odpowiedni dobór metody oceny efektu w stosunku do badanego zagadnienia. Dodatkowym wymogiem jest określenie horyzontu czasowego oceny skutków, jako że moment obserwacji może wpływać na końcowe efekty.

Wynikiem analizy są współczynniki koszt-efektywność (*ang. cost-effectiveness ratio, CER*), obliczone dla poszczególnych programów objętych badaniem. Powyższy współczynnik definiowany jest jako stosunek poniesionych kosztów do uzyskiwanych efektów (Równanie 16).

**Równanie 16** Współczynnik koszt-efektywność (CER).

$$\text{CER} = \frac{\text{koszt schematu terapeutycznego}}{\text{efekt schematu terapeutycznego}}$$

Jednocześnie jednak podejmowanie decyzji o charakterze ekonomicznym i terapeutycznym, na podstawie współczynników CER nie jest praktyką prawidłową. Efektem uzyskania wartości CER jest informacja na temat kosztu jednostki wyniku (efektu), podczas gdy informacją, która może być pomocna przy wspomnianych decyzjach, jest koszt uzyskania dodatkowej jednostki efektu. Wartość ta określana jest jako inkrementalny współczynnik koszt-efektywność (*ang. incremental cost-effectiveness ratio, ICER*) i definiowana jako stosunek różnicy kosztów między porównywanymi programami do różnicy wyników porównywanych programów (Równanie 17).

**Równanie 17** Inkrementalny współczynnik koszt-efektywność (ICER).

$$\text{ICER} = \frac{\text{różnica kosztów pomiędzy porównywanymi schematami}}{\text{różnica efektów pomiędzy porównywanymi schematami}}$$

Analiza inkrementalna przeprowadzana jest w każdym przypadku kiedy występuje dysproporcja między kosztami i wynikami badanych opcji terapeutycznych:

- Program  $A_{\text{koszt}} > \text{Program } B_{\text{koszt}}$  oraz Program  $A_{\text{efekt}} < \text{Program } A_{\text{efekt}}$
- Program  $A_{\text{koszt}} < \text{Program } B_{\text{koszt}}$  oraz Program  $A_{\text{efekt}} > \text{Program } A_{\text{efekt}}$

Kolejny etap to analiza wrażliwości. Jej celem jest określenie wpływu ewentualnych zmian przyjętych założeń na wyniki całego badania, stąd określenie jej mianem analizy „co jeśli” (*ang. „what if” analysis*). Prawidłowo przeprowadzona pozwala zredukować ryzyko niepewności uzyskiwanych wyników końcowych. Analiza wrażliwości powinna być przeprowadzana według podanego schematu [96]:

- a) identyfikacja elementów niepewnych, czyli takich, których zmiana może w sposób istotny wpłynąć na wynik końcowy

Teoretycznie wszystkie zmienne, na podstawie których wykonano analizę powinny być poddane analizie wrażliwości. Obecnie najczęściej analizie poddawane są najważniejsze dla przeprowadzanego badania dane wybrane przez autora doświadczenia [130].

b) określenie badanego zakresu zmienności elementów krytycznych

Fakt doboru *a-priori* zarówno czynników poddawanych analizie jak i zakresu zmienności jest krytykowany jako poważne ograniczenie analizy wrażliwości. Wiarygodny badany zakres zmienności powinien być zbudowany na podstawie rzetelnych danych. Przyjmuje się, że dostateczną pewnością cechują się przedziały oparte o opinie ekspertów, przegląd literatury bądź przedział ufności wokół średniej dla danych losowych.

c) rekalkulacja wyników w przyjętym zakresie określonych wcześniej parametrów

Końcowy etap to ponowne wyliczenie i porównanie współczynników efektywności kosztów. Uzyskiwane różnice pozwalają ocenić w jakim stopniu poszczególne składowe analizy wpływają na jej wynik. W przypadku uzyskanego braku wrażliwości, przyjęte parametry wstępne mogą zostać ocenione jako właściwe. Wynik przeciwny pozwala zwrócić uwagę na elementy, które w przypadku podejmowanych na podstawie wyników analizy decyzji o charakterze finansowym, mogą powodować zmianę współczynników.

W praktyce wykorzystywanych jest kilka podstawowych rodzajów analizy wrażliwości [96]:

- prosta analiza wrażliwości (analiza jednokierunkowa) – schemat jej przeprowadzania opiera się na badaniu wpływu zmiany kolejnych, pojedynczych parametrów na wynik końcowy
- prosta analiza wrażliwości (analiza wielokierunkowa) – umożliwia określenie wpływu kilku zmiennych równocześnie (w przyjętym zakresie) co w większym stopniu odpowiada warunkom rzeczywistym; jako, że możliwa liczba kombinacji rośnie wykładniczo w stosunku do liczby czynników poddawanych analizie, konieczne

jest zastosowanie modeli eksperymentalnych przedstawiających możliwe scenariusze; jedną z metod uniknięcia nadmiernej ilości pracy o charakterze obliczeniowym jest zastosowanie analizy wartości ekstremalnych (skrajnych) [96]; skonstruowanie tzw. scenariuszy – podstawowego, najbardziej optymistycznego oraz najbardziej pesymistycznego i rekalkulacja wyników w oparciu o nie pozwala ocenić wrażliwość poszczególnych składowych

- analiza progowa – polega na określeniu krytycznych wartości parametrów, dla których dochodzi do odwrócenia końcowych wyników analizy; pozwala to na określenie kombinacji parametrów, dla której dochodzi do przekroczenia progu akceptacji wybranego programu
- analiza probabilistyczna – uwzględnia względne prawdopodobieństwo wystąpienia poszczególnych wartości (w przeciwieństwie do analizy wartości skrajnych); wykorzystywane są różne możliwości modelowania (np. metoda Monte-Carlo)

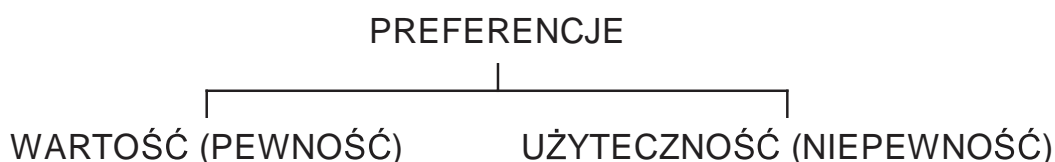
Celem przeprowadzania analizy koszt-efektywność jest możliwość wyboru najbardziej efektywnej drogi terapeutycznej, z możliwych do zastosowania przy określonych (ograniczonych) zasobach. Jej wyniki pomagają podejmować decyzje o alokacji środków finansowych, dlatego ich odbiorcami mogą być zarówno menedżerowie zarządzający służbą zdrowia, np. podejmujący decyzje o refundacji leków jak i praktykujący lekarze.

#### 2.2.4.4 Analiza użyteczności kosztów (koszt-użyteczność)

Analiza użyteczności kosztów jest specyficzną formą oceny, w której jednostka wyniku definiowana jest jako użyteczność, czyli zdolność danego zasobu do zaspokojenia określonych potrzeb ludzkich. Analiza CUA traktowana jest niekiedy jako specjalny rodzaj analizy efektywności kosztów. Może być również traktowana jako jej rozszerzenie, jako że umożliwia porównanie technologii medycznych w sytuacji, kiedy zastosowanie analizy CEA nie jest możliwe [38].

Użyteczność jest terminem przedstawionym po raz pierwszy przez von Neumanna i Morgensterna w roku 1944 jako część teorii gier. Ograniczony założeniami (m.in. przechodniość, niezależność i ciągłość preferencji) model von Neumanna-Morgensterna opisuje matematycznie w jaki sposób jednostka racjonalna powinna podejmować decyzje w sytuacji ryzyka. Jako, że klasyczna teoria przedstawionego modelu zakłada podejmowanie decyzji obarczonych niepewnością (użyteczność) konieczne jest jej zróżnicowanie w stosunku do preferencji i wartości (podejmowanie decyzji w warunkach pewności) oraz przedstawienie odpowiednich narzędzi pomiaru.

**Schemat 15** Zależność między preferencją oraz wartością i użytecznością.



Jak wynika z powyższego schematu pojęcie preferencji jest pojęciem szerszym w stosunku do użyteczności i wartości, które są jej podtypami. W związku z tym inne są metody ich pomiaru i zastosowanie.

**Tabela 2** Metody pomiaru preferencji [38].

Metoda odpowiedzi	Metoda formułowania pytań	
	Wartość (pewność)	Użyteczność (niepewność)
Skala	<ul style="list-style-type: none"> <li>– skala punktowa</li> <li>– skala kategorii</li> <li>– skala analogowa</li> <li>– skala współczynnikowa</li> </ul>	
Wybór	<ul style="list-style-type: none"> <li>– handlowanie czasem</li> <li>– porównanie par</li> <li>– ocena równoważności</li> <li>– handlowanie liczbą osób</li> </ul>	<ul style="list-style-type: none"> <li>– loteria</li> </ul>

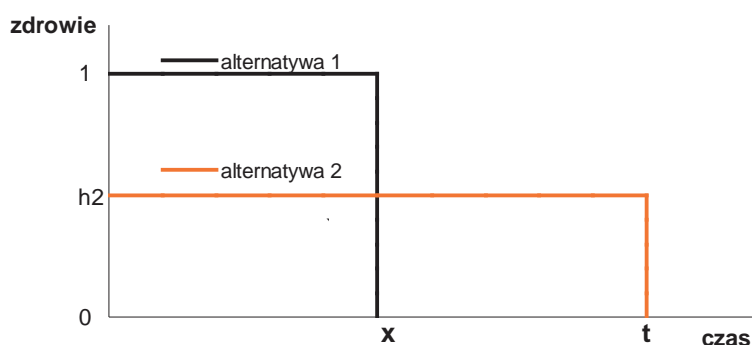
Najczęściej spotykane w piśmiennictwie metody pomiaru preferencji to skala wartościowania, handlowanie czasem, handlowanie liczbą osób oraz loteria.

Wykorzystanie skali wartościowania (*ang. rating scale*) możliwe jest dzięki założeniu, że ludzkie, subiektywne oceny można kwantyfikować. W badaniach wykorzystuje się zarówno skalę interwałową jak i porządkową. Przy zastosowaniu skali interwałowej badany po uszeregowaniu stanów zdrowia od najmniej do najbardziej pożądanego odcina wyniki na linii prostej. Zasadniczą rzeczą jest aby odległości (interwały) pomiędzy zaznaczonymi punktami odpowiadały różnicy w ocenie, co umożliwi stworzenie interwałowej skali preferencji. Stosunek bezwzględnych wartości umieszczonych na skali interwałowej nie ma znaczenia, dlatego na ich podstawie niemożliwe jest dokonanie wartościowania.

Skala porządkowa stosowana jest najczęściej w przypadku chorób przewlekłych. Badany wskazuje stan najlepszy (najbardziej preferowany, oznaczany jako 1) oraz najgorszy (najmniej preferowany, najczęściej definiowany jako „zgon” choć niekoniecznie, oznaczany cyfrą 0) stan zdrowia. Następnie stany przejściowe umieszczane są na linii prostej, w taki sposób aby odległość między nimi była proporcjonalna do różnic preferencji chorego. Wykorzystując skalę porządkową w przypadku chorób przejściowych jako koniec przedziału umieszczany jest najgorszy możliwy stan choroby okresowej. Wyniki uzyskane z wykorzystaniem skali porządkowej mogą zostać przetransponowane na punktację uzyskaną inną metodą (np. handlowania czasem).

Metoda handlowania czasem (*ang. time trade-off*) oparta jest na określeniu czasu, o który badany byłby gotowy skrócić życie w zamian za poprawę stanu zdrowia.

**Schemat 16** Graficzny opis metody handlowania czasem dla chorób przewlekłych [38,96].



Schemat 16 przedstawia graficzną reprezentację oceny preferencji za pomocą metody handlowania czasem dla chorób przewlekłych. Na osi zdrowia wartość 1 oznacza pełne zdrowie, podczas gdy wartość 0 koduje zgon. Porównano dwa hipotetyczne stany zdrowia uzyskane dla alternatywy 1 (pełnia zdrowia przez krótszy czas) oraz alternatywy 2 (dłuższy okres przeżycia w gorszym stanie zdrowia), które przedstawia się badanemu. Wartość czasu  $x$  zmienia się do momentu, w którym obie alternatywy są dla badanego równoważne. Poszukiwany wynik preferencji obliczany jest zgodnie z wzorem, który przedstawia Równanie 18.

**Równanie 18** Schemat obliczenia wartości preferencji metodą handlowania czasu w przypadku chorób przewlekłych.

$$\text{wartość preferencji } h_2 = \frac{x}{t}$$

gdzie:

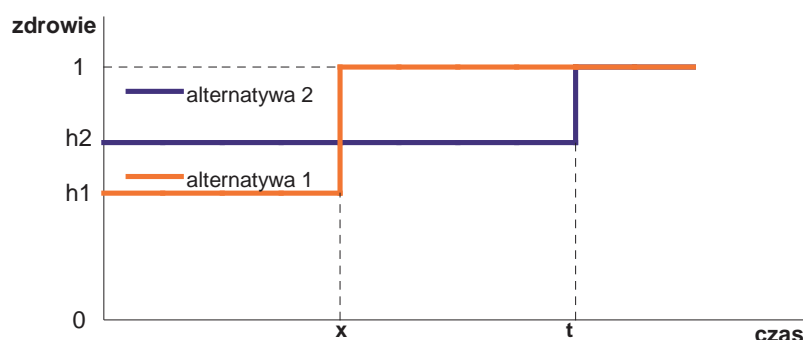
$x$  – okres trwania stanu zdrowia uzyskanego dzięki zastosowaniu alternatywy 2

$t$  - okres trwania stanu zdrowia uzyskanego dzięki zastosowaniu alternatywy 1

Zastosowanie metody handlowania czasem do oceny preferencji jest możliwe również w przypadku chorób przejściowych [38,96].



**Schemat 17** Graficzny opis metody handlowania czasem dla chorób przejściowych.



Wartość czasu  $x$  zmienia się, aż do momentu kiedy obie alternatywy stają się równoważne. Poszukiwany wynik preferencji obliczany jest z Równanie 19.

**Równanie 19** Schemat obliczenia wartości preferencji metodą handlowania czasem w przypadku chorób przejściowych.

$$\text{wartość preferencji } h_2 = \frac{1 - (1 - h_1)x}{t}$$

gdzie:

$x$  – okres trwania stanu zdrowia uzyskanego dzięki zastosowaniu alternatywy 2

$t$  - okres trwania stanu zdrowia uzyskanego dzięki zastosowaniu alternatywy 1

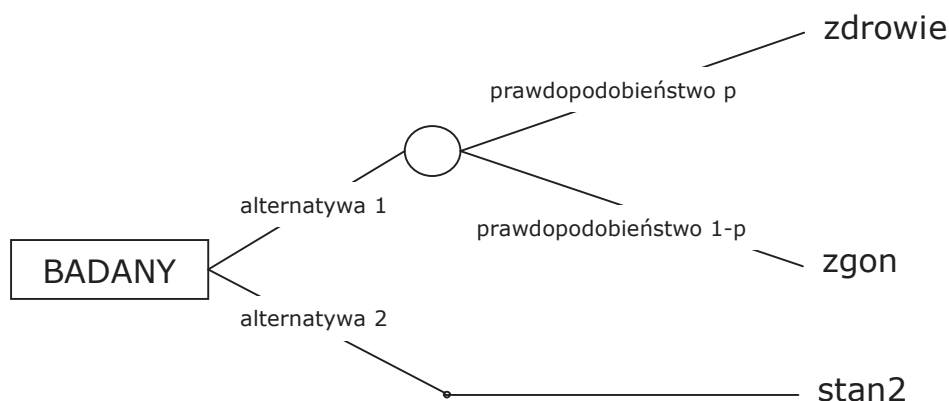
$h_1$  – stan zdrowia w czasie  $x$  (najgorszy możliwy stan zdrowia, w każdym przypadku dla którego istnieje stan zdrowia gorszy niż  $h_2$ )

Opisana powyżej metoda konwencjonalnego handlowania czasem ma swoje rozwinięcia [38]. Dwie inne możliwości to roczne handlowanie czasem, gdzie badany dokonuje wymiany wobec niechcianej rekonwalescencji oraz dzienne handlowanie czasem, gdzie podstawą wymiany jest niepożądany sen.

Handlowanie ilością osób to metoda analogiczna do metody handlowania czasem, w której źródłem oceny preferencji jest określenie przez badanego ilości osób w określonym stanie zdrowia  $a$ , które należy wyleczyć w ciągu roku aby zrównoważyć uratowanie życia jednej osobie w ciągu roku. Ze względu na społeczny wydzźwięk przedstawiona metoda stosowana jest przez ubezpieczycieli zdrowotnych.

Loteria to klasyczna metoda pomiaru preferencji kardynalnych. Jej wynikiem jest użyteczność, jako że zawiera element niepewności. Wykorzystywana jest zarówno do oceny użyteczności poszczególnych alternatyw w chorobach przewlekłych jak i stanów przejściowych. Schemat 18 przedstawia metodykę oceny. W pierwszym etapie choremu prezentowane są dwie alternatywy: leczenie, które z określonym prawdopodobieństwem  $p$  może doprowadzić do osiągnięcia stanu doskonałego zdrowia, w którym pacjent żyje dodatkową liczbę  $t$  lat oraz z prawdopodobieństwem  $1-p$  do zgonu. Alternatywą drugą jest stan przewlekły, trwający  $t$  lat. Podczas oceny zmianie ulega prawdopodobieństwo  $p$ , aż do momentu, w którym obydwie alternatywy – w ocenie pacjenta – są sobie równe. Wtedy wynik preferencji  $h_1$  równy jest  $p$  (Równanie 20).

**Schemat 18** Ocena preferencji dla stanu przewlekłego przy zastosowaniu metody loterii.



**Równanie 20** Schemat obliczenia wartości preferencji metodą loterii w przypadku chorób przewlekłych.

$$\text{wartość preferencji } h_2 = p$$

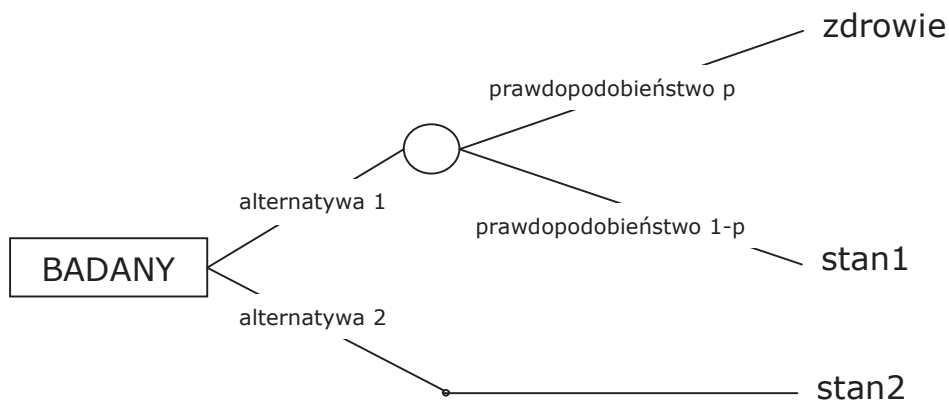
gdzie:

$p$  – prawdopodobieństwo, przy którym wyrównywane są preferencje pacjenta

W przypadku stosowania loterii do oceny preferencji pacjenta w krótkotrwałym stanie chorobowym, dolną granicą oceny nie jest zgon lecz

najgorszy możliwy stan przejściowy. Konieczne do spełnienia założenie obejmuje równość czasów trwania  $t$  wszystkich stanów. Metodykę oceny przedstawia Schemat 19.

**Schemat 19** Ocena preferencji dla stanu przejściowego przy zastosowaniu metody loterii.



**Równanie 21** Schemat obliczenia wartości preferencji metodą loterii w przypadku chorób przejściowych.

$$\text{wartość preferencji } h_2 = p + (1-p)h_1$$

gdzie:

$p$  – prawdopodobieństwo, przy którym wyrównywane są preferencje pacjenta

$h_1$  – najgorszy możliwy stan zdrowia

Wartość  $h_1$  otrzymywana jest w wyniku zastosowania metody loterii dla chorób przewlekłych, gdzie najgorszy możliwy stan to zgon, a skala wartości  $p$  mieści się w przedziale  $\langle 0;1 \rangle$ . Poszukiwana wartość  $h_1$  jest podstawiana w miejsce wartości  $h_2$  w Równanie 20.

Wyniki pomiaru preferencji uzyskane różnymi metodami najczęściej różnią się od siebie [38,96]. Przyczyną jest różne ujęcie zagadnienia ryzyka oraz różna metodyka oceny. Najczęściej poleca się stosować pomiar preferencji wyrażonej w postaci użyteczności, jako że przyszłe rezultaty zdrowotne obarczone są niepewnością. Niemniej jednak wszystkie metody wymienione w Tabeli 2 są wykorzystywane w praktyce.

Przedstawione powyżej sposoby pomiaru preferencji są czasochłonne i kłopotliwe w zastosowaniu praktycznym. W związku z tym w badaniach wykorzystuje się obecnie najczęściej gotowe wieloczynnikowe systemy klasyfikacji stanu zdrowia, co wymagało rozwinięcia klasycznej teorii użyteczności von Neumanna-Morgensterna, do teorii wieloczynnikowej (postawienie dodatkowego założenia dotyczącego niezależności użyteczności, wyrażanej jako niezależność pierwszorzędowa, multiplikatywna lub addytywna) [38,96].

Trzy najczęściej wykorzystywane systemy to: Skala Jakości Dobrostanu (*ang. Quality of Well-Being, QWB*), Indeks Użyteczności Zdrowia (*ang. Health Utilities Index, HUI*) oraz EuroQol.

System QWB klasyfikuje pacjentów względem czterech cech: zdolność poruszania się, aktywność fizyczna, aktywność społeczna oraz złożoną jakość objaw-zaburzenie. Obliczenie parametrów skali oparte jest na wartościach uzyskanych w wyniku doświadczenia przeprowadzonego na losowej grupie osób, które umiejscawiają różne stany zdrowia na skali od 0 (zgon) do 1 (pełne zdrowie).

Metoda HUI dzieli się na trzy dostępne podsystemy. Punktacja opiera się na preferencjach grup populacji (skala interwałowa) i obejmuje cztery cechy: funkcjonowanie fizyczne (poruszanie się i aktywność fizyczna), funkcjonowanie społeczne (uwzględnia codzienne czynności) funkcjonowanie społeczno-emocjonalne (aktywność społeczna oraz emocjonalny dobrostan) oraz zaburzenia zdrowotne. Wskaźniki preferencji zostały obliczone przy wykorzystaniu wyników uzyskanych na losowej grupie metodą handlowania czasem.

Skala EuroQol zawiera pięć elementów: zdolność poruszania się czynności codzienne, aktywności główne, ból/dyskomfort oraz lęk/depresja, z których każdy może być przedstawiany na trzech poziomach: brak zaburzeń, niewielkie zaburzenia, poważne zaburzenia. Preferencja została zmierzona metodą handlowania czasem na losowej grupie osób.

W analizie użyteczności kosztów wyniki stosowanej procedury medycznej wyrażane są najczęściej w postaci QALY czyli roku życia skorygowanego o jakość, a wynikiem ostatecznym analizy jest koszt przypadający na zyskaną jednostkę QALY (inkrementalny współczynnik użyteczności kosztów). Zaletą QALY jest możliwość jednoczesnego przedstawienia korzyści będących pochodną ograniczenia chorobowości (zysk jakościowy) oraz śmiertelności (zysk ilościowy) [38,96]. Wartość QALY może zostać obliczona za pomocą wszystkich wymienionych powyżej metod oceny preferencji na dwóch drogach: sumowanie poszczególnych składowych oraz pomiar użyteczności profilu zdrowia. Aby móc zaklasyfikować obliczoną wartość QALY jako użyteczność musi ona spełniać wszystkie założenia jej dotyczące.

Alternatywnymi dla QALY jednostkami są: równoważnik roku życia przeżytego w pełnym zdrowiu (*ang. Healthy-Year Equivalent, HYE*) oraz równoważnik jednej uratowanej od śmierci młodej osoby (*ang. Saved-Young-Life Equivalent, SAVE*) [38,96].

W metodzie HYE ocenia się preferencje dotyczące wszystkich stanów zdrowia w całym badanym okresie czyli tzw. profil zdrowia. Wykorzystuje się dwustopniową metodę loterii oceniając zarówno użyteczność danego profilu zdrowia jak i czas trwania w pełnym zdrowiu. Pomiar HYE jest metodologicznie oraz praktycznie trudniejszy niż pomiar QALY.

Metoda SAVE opiera się na handlowaniu liczbą osób, gdzie badanemu zadawane jest pytanie dotyczące ilości osób w danym stanie zdrowia, które należałoby wyleczyć aby zrównoważyć uratowanie od śmierci jedną młodą osobę. Wykorzystując SAVE otrzymuje się wyniki, które obrazują społeczny, a nie indywidualny punkt widzenia (gra się życiem innych nie własnym, stąd większa tendencja do zawyżania ryzyka).

Analiza użyteczności kosztów jest metodą przydatną w pewnych specyficznych sytuacjach, kiedy na pierwszy plan wysuwają się zagadnienia związane z jakością życia. Dodatkowo wolna jest od niektórych ograniczeń innych typów analiz (np. możliwość porównywania programów o różnych wynikach w przeciwieństwie do analizy efektywności

kosztów lub w razie niemożności przedstawienia wyników w postaci pieniężnej jak w analizie wydajności kosztów).

#### 2.2.4.5 Analiza wydajności kosztów (koszt-korzyści)

Analiza wydajności kosztów (CBA) jest jednym z typów pełnej analizy farmakoekonomicznej, jako że wymaga oceny zarówno kosztów jak i efektów stosowanych procedur [38,96]. Założenia tego typu oceny farmakoekonomicznej wywodzą się z teorii ekonomii dobrobytu, której podstawy teoretyczne zostały sformułowane przez Wilfredo Pareto. Podstawowym zagadnieniem opisywanym przez ten dział ekonomii jest kwestia odpowiedzi na pytanie jak oceniać czy społeczeństwo jako całość zyskuje na prowadzonej procedurze/polityce (jest to więc podejście wartościujące, a nie tylko opisujące sytuację).

Skutki zdrowotne będące wynikiem zastosowania danej procedury medycznej wyrażane są w postaci wartości pieniężnej. Korzyści wynikające z zastosowania analizy wydajności kosztów w porównaniu z powyżej przedstawionymi typami analiz wynikają z możliwości pominięcia jednego z założeń charakterystycznych zarówno dla analizy efektywności kosztów jak użyteczności kosztów, a dotyczącego konieczności przedstawienia wszystkich możliwych alternatyw ocenianej procedury [38,96,103]. Podobne założenie nie jest konieczne w przypadku analizy CBA.

Podejmowanie decyzji finansowych na podstawie analiz farmakoekonomicznych wymaga przedstawienia wyniku w postaci monetarnej. W przypadku analizy CEA oraz CUA konieczne jest odwołanie do pewnych wartości względnych (gotowe tabele, wartości progowe). W przypadku analizy CBA nie jest to wymagane, co w sytuacji konieczności podjęcia pojedynczej decyzji powoduje, że jest ona bardziej wartościowa w stosunku do analiz CEA i CUA.

Podstawowym zagadnieniem odróżniającym metodologicznie opisywaną analizę od przedstawionych powyżej, jest sposób oszacowania

efektów zdrowotnych, poprzez nadanie im wartości pieniężnych. W literaturze problemu wyróżnia się trzy podstawowe metody oceny:

a) metoda kapitału ludzkiego

Metoda kapitału ludzkiego wywodzi się bezpośrednio z ekonomii dobrobytu. Środki przeznaczone na uzyskiwanie danego efektu medycznego traktowane są jako inwestycja w kapitał ludzki. W praktyce oznacza to, że opłacalność poniesionych nakładów określa się jako przywróconą lub zwiększoną wydajność danej osoby na rynku pracy, a więc czasowi wolnemu od choroby lub jej objawów została nadana pewna wartość pieniężna. W tym celu wykorzystywana jest wartość rynkowa pracy (stawka płac), a wartość badanej procedury wyrażona jest w postaci przyszłych zarobków. Metoda kapitału ludzkiego wykorzystywana jest jako jedyne źródło oceny zmiany stanu zdrowia lub jako ocena częściowa tylko dla oceny wartości zmian produktywności jednostki.

Metoda ta jest krytykowana ze względu na niedoskonałości ujawniające się w trakcie przeprowadzania pomiarów. Podstawowa trudność dotyczy dokładności pomiarów, a wynikająca z niedoskonałości rynku pracy (zróżnicowanie stawek nie tylko w zależności od wydajności). Dodatkowy kłopot związany jest z oceną zasobów nierynkowych (np. wartościowanie pracy osób nie wykonujących pracy zarobkowej czego przykładem mogą być gospodynie domowe).

b) badanie ustalonych preferencji

Metoda ma na celu określenie preferencji dotyczących zamiany stanu zdrowia na wynagrodzenie. Przykładem może być wybór między pracą w warunkach o zwiększonym ryzyku zmiany stanu zdrowia (wypadku przy pracy, zgonu) przy jednoczesnym wyższym wynagrodzeniu oraz pracą bezpieczniejszą ale gorzej płatną. Metoda ta opiera się na rzeczywistym wyborze między dostępnymi opcjami, nie bazując na ustaleniach teoretycznych. Podobnie jak w metodzie kapitału ludzkiego ograniczeniem jest niedoskonałość rynku pracy oraz wpływ czynników zakłócających zależność poziom ryzyka – wynagrodzenie.

Alternatywą jest metoda analizy podjętych decyzji ujawniająca społeczną ocenę wartości poszczególnych efektów zdrowotnych.

c) badanie metodą oceny warunkowej

Badania metodą oceny warunkowej oparte są na ankietach, w których respondenci zaznaczają maksymalną cenę, którą są gotowi zapłacić za określoną procedurę lub korzyść kliniczną [38]. Jednym z założeń koniecznych do przeprowadzenie podobnego badania jest przyjęcie istnienia idealnie wolnego rynku usług medycznych. Badanie wykorzystujące metodę określaną jako „gotowość do płacenia” (*ang. willigness-to-pay, WTP*) może zostać wykonane w warunkach pewności, przy założeniu potencjalnych (niepewnych) korzyści zdrowotnych (*ex-post*), a także w warunkach niepewności zarówno co do schorzenia jak i efektów zdrowotnych (*ex-ante*). Umożliwia to dokonanie pomiaru użyteczności zarówno z punktu widzenia korzyści pojedynczego pacjenta, korzyści prywatnych jak i społeczeństwa. Korzyść pacjenta uzyskiwana jest przez obliczenie średniej WTP w populacji i przeliczeniu jej na jednego pacjenta. Korzyści prywatne uzyskiwane są na dzięki pomnożeniu średniej WTP uzyskanej z pytań *ex-ante* przez liczbę osób z grupy ryzyka zachorowania na daną jednostkę chorobową. Aby obliczyć korzyści społeczne konieczne jest skorygowanie średniej WTP o dystrybucję dochodów i pomnożenie jej przez liczbę osób populacji.

Jako, że przedstawienie korzyści zdrowotnych w wartościach monetarnych – mimo rozbudowanej metodyki – jest trudne, metoda CBA wykorzystywana jest stosunkowo rzadko. Analiza wydajności kosztów może być szczególnie przydatna w przypadku podejmowania decyzji o alokacji zasobów, umożliwia bowiem ocenę porównawczą programów o drastycznie różnych wynikach. Ostatecznym wyznacznikiem oceny jest zysk netto osiągany przy ich pomocy. Dzięki temu dane zasoby finansowe mogą być wykorzystane do uzyskania maksymalnych efektów, bądź dane efekty można uzyskać angażując minimum środków.



#### 2.2.4.6 Analiza koszt-konsekwencje

Analiza konsekwencji kosztów jest najprostszą formalnie i jednocześnie najwszechstronniejszą prezentacją informacji dotyczącej stosowanych procedur medycznych [96]. W najpełniejszej formie analiza CCA powinna dostarczać informacji na temat wszystkich dostępnych, alternatywnych technologii medycznych oraz jednocześnie wszystkie możliwości wyników terapeutycznych. Uzyskana lista kosztów i konsekwencji może sama w sobie stanowić narzędzie umożliwiające porównanie jak i być podstawą innych analiz farmakoekonomicznych, jako że umożliwia całościowe spojrzenie i wybór informacji koniecznych w poszczególnych typach analiz.

Najważniejszą cechą analizy CCA jest brak gotowego systemu wartościowania przedstawianych informacji. Użytkownik korzystający z jej wyników samodzielnie przyporządkowuje odpowiednie cechy do grupy korzyści lub wad na podstawie własnego systemu oceny.

Zakres informacji zbieranych w trakcie przeprowadzania analizy konsekwencji kosztów pozwala najbardziej obiektywnie ocenić całość zagadnienia związanego z daną technologią medyczną (np. farmakoterapią określonego schorzenia) ale jednocześnie może stanowić ograniczenie ze względu na trudności w bezpośrednim porównaniu różnych danych.

#### 2.2.5. Modelowanie w farmakoekonomicie

Jak przedstawiono powyżej (Rozdział 2.2) analiza farmakoekonomiczna obejmuje zestawienie konsekwencji i kosztów leczenia, co z kolei daje możliwość specyficznego porównania stosowanych procedur. Konieczność zgrupowania dwóch rodzajów danych, może limitować dostępność tej grupy metod, stąd wykorzystanie technik modelowania matematycznego.

Wydawać by się mogło, że w przypadku ograniczeń metodologicznych utrudniających zastosowanie klasycznych technik modelowania, narzędzia inteligencji obliczeniowej stanowić mogą ich

cenne uzupełnienie. Niestety mimo kilkuletnich poszukiwań, autorowi nie udało się znaleźć publikacji lub doniesienia o charakterze naukowym, omawiającego bezpośredniego zastosowania powyższych metod w farmakoekonomice. Jednocześnie jednak prowadzone są doświadczenia z zastosowaniem sztucznych sieci neuronowych jako składników systemów ekspertowych konstruowanych w celu optymalizacji wydatków w opiece zdrowotnej. Walczak oraz współautorzy [132] testowali możliwość wykorzystania sieci neuronowych, do ograniczania kosztów związanych z terapią krwiozastępczą w trakcie zabiegów chirurgicznych. Celem pracy była optymalizacja ilości preparatów krwiopochodnych, zamawianych każdorazowo przed zabiegiem chirurgicznym. Cytowane przez autorów publikacje informowały o ich nadmiernym wykorzystaniu, co zarówno ze względu na generowane koszty jak i niepoliczalną wartość preparatów krwiopochodnych wymaga wyeliminowania. Klasyczne procedury liczenia zapotrzebowania, oparte na technikach statystycznych zestawione zostały z modelem neuronowym. W trakcie doświadczeń wykorzystano sieci o radialnych funkcjach bazowych - RBF (*ang. Radial Basis Function*). Dane pierwotne zawierały wyniki niektórych badań biochemicznych (hematokryt, czas protrombinowy, częściowy czas tromboplastyny, liczba płytek krwi), informacje na temat wieku, klasy do jakiej zaliczono pacjenta wg wytycznych ASA (*ang. American Society of Anesthesiologists*) oraz ew. obecność dodatkowych schorzeń układów krążenia i oddechowego. Podczas doświadczeń testowano różny rozmiar wektora wejściowego, w tym także zredukowany do sześciu składników, poprzez eliminację danych skorelowanych. Wyjście miało charakter ciągły i oznaczało zapotrzebowanie na preparat krwiopochodny wyrażany w jednostkach ilościowych oraz jako współczynnik C/T (*ang. crossmatched/transfused*) czyli iloraz ilości krwi przygotowanej do przetoczenia (przebadanej) do rzeczywiście przetoczonej. Trafność predykcji oceniana była za pomocą testu t-Studenta oraz z wykorzystaniem RMSE (*ang. Root Mean Square Error*). Jako końcowy wniosek autorzy przedstawiają wyliczenia dotyczące ew. oszczędności płatnika (szpital, pacjent) polegających na obniżeniu

kosztów zarządzania zasobami krwi wykorzystywanej w trakcie przykładowego zabiegu. Oszczędności obejmują zarówno pracę techników medycznych, przygotowujących krew przed operacją jak i samych zasobów krwi. Z kalkulacji wynika, że potencjalna oszczędność wynikająca z zastosowania opisanej procedury może wynieść ponad 3 tys. dolarów na każde 50 przeprowadzonych wzorcowych zabiegów, których dane wykorzystano w trakcie doświadczeń.

Podobnie Walczak i współautorzy [131] testowali możliwość zastosowania sieci neuronowych zaliczanych do różnych kategorii jako narzędzi optymalizacji wykorzystania zasobów szpitalnych. Problem dotyczył modelowania długości pobytu chorego na oddziale intensywnej terapii co jednocześnie pozwala zbudować plan wykorzystania łóżek szpitalnych tego specyficznego oddziału. Wśród testowanych metod znalazły się klasyczne perceptrony wielowarstwowe, sieci typu RBF oraz samoorganizujące sieci rozmyte ARTMAP. Wyjście miało charakter ciągły, a skuteczność systemów oceniano wykorzystując wartość RMSE. Uzyskanie skutecznego narzędzia pozwalającego planować zużycie posiadanych zasobów stanowi jednocześnie krok w kierunku ewentualnych oszczędności, których uzyskanie wynika z możliwości optymalnej alokacji środków.

## 2.3 Epidemiologia

### 2.3.1. Wstęp

Epidemiologia jest uważana za jedną z podstawowych nauk z zakresu zdrowia publicznego [105]. Pierwotnie pojmowana epidemiologia obejmowała opis i co za tym idzie analizę zjawisk związanych z rozprzestrzenianiem się chorób zakaźnych, jednak rozwój nauki i ewolucja społeczeństw znacznie poszerzyły zakres jej zainteresowań, obejmując także problemy związane z budową modeli epidemiologicznych. Oznacza to, że obecnie epidemiologia oprócz podejścia klasycznego oferuje również metody wnioskowania oparte na testowaniu hipotez naukowych

pozwalających m.in. na kontrolę zachorowalności i umieralności, co z kolei wpływa na stan zdrowia społeczeństw.

### 2.3.2. Modelowanie w epidemiologii

Modelowanie epidemiologiczne ma długą historię i bogatą literaturę [12,15,44,49,56,59,78,85,136]. Jego ewolucja przebiega w sposób równoległy, w stosunku do rozwoju metod modelowania. Powyższa prawidłowość obejmuje zarówno modele opisowe jak i predykcyjne, i dotyczy wszystkich możliwości oferowanych przez współczesną statystykę. Stosowane w epidemiologii techniki modelowania obejmują większość znanych metod klasycznej statystyki (techniki regresji liniowej i nieliniowych, metoda drzew decyzyjnych, model Markova i inne).

Podział możliwości wykorzystania sztucznych sieci neuronowych w epidemiologii przeprowadzany jest ze względu na kierunki zastosowań. Grupa pierwsza obejmuje zagadnienia klasyfikowane następnie pod kątem modelowanych wartości (modelowanie specyficznych parametrów epidemiologicznych) oraz ze względu na rodzaj wykorzystywanych danych (proste dane demograficzne lub antropometryczne np. w przypadku modeli neuronowych do wstępnego skriningu chorób lub ich powikłań).

#### 2.3.2.1 Modele skriningowe

Mohamed i współautorzy zaproponowali zastosowanie modeli neuronowych jako narzędzi do predykcji wartości BMD (*ang. Bone Mineral Density*) [89]. Jest to propozycja alternatywna w stosunku do klasycznej metody pomiaru absorpcji promieniowania rentgenowskiego (*ang. dual X-ray absorptiometry - DXA*), której wykorzystanie na skalę masową ograniczają kosztowność oraz słaba dostępność. Określenie wartości BMD jest pomocne przy diagnozowaniu zmian o charakterze osteopenii i osteoporozy. Ich wczesne zdiagnozowanie może pozwolić uniknąć powikłań, które najczęściej manifestują się w postaci złamań i co za tym idzie cierpień pacjentów oraz drastycznie wysokich kosztów związanych z ich leczeniem.

Dane wejściowe stanowiły proste informacje antropometryczne – płeć, wiek, wartość BMI, współczynnik talia-biodra, wartości pomiaru fałdów skórnych oraz wartości masy niektórych tkanek (np. tkanka tłuszczowa). Bazę danych stanowiły informacje opisujące 1154 pacjentów posegregowanych w grupy w zależności od obecności schorzenia lub jego braku. W trakcie badań wykorzystano sieci neuronowe uczone przy pomocy algorytmu wstecznej propagacji błędu modyfikacją APROP oraz sigmoidalną funkcję aktywacji. Jako narzędzia walidacyjnego użyto metody *leave-one-out*. Zarówno modele otrzymane na pełnym wektorze wejściowym (7 wejść) jak i po jego redukcji (4 wejścia) były w stanie podać wyniki nie różniące się istotnie statystycznie od wyników otrzymanych przy pomocy metody DXA.

Wśród ciekawszych zastosowań sztucznych sieci neuronowych w szeroko pojętej epidemiologii wymienić można ich zastosowanie w modelowaniu przebiegu choroby Alzheimera na podstawie prostych informacji demograficznych [124]. Informacją modelowaną była wartość MMSE (*ang. Mini-Mental State Examination*) – wskaźnik opisujący zaawansowanie demencji na podstawie oceny zdolności poznawczych. Wektor wejściowy stanowiły dane kodujące płeć, wykształcenie, wstępną diagnozę, zaawansowanie choroby w momencie diagnozy w skali 1-5 przy użyciu współczynnika CDR (*ang. Clinical Dementia Rating Scale*), współczynnik progresji choroby. Dodatkowo jedno z wejść podawało informację dotyczącą statusu społecznego (czy pacjent mieszka sam). Zadaniem porównywanych modeli była ocena zmian współczynnika MMSE w czasie. Autorzy przeprowadzili test porównawczy modelu liniowego, klasycznej sieci neuronowej opartej o algorytm wstecznej propagacji błędu oraz zaproponowanego przez siebie rozwiązania opartego o sieć neuronowa określanego jako *mixed effects neural network*. Jego specyfika polega na możliwości bezpośredniego wykorzystania informacji z obserwacji rozciągniętych w czasie. Autorzy wykorzystywali technikę 10-cio krotnej krzyżowej walidacji. Wartość błędu MSE (jedna z wykorzystywanych miar błędów) dla najlepszego otrzymanego modelu

wynosiła 0,37. Uzyskane wyniki pozwoliły wyciągnąć wniosek o potencjalnej przydatności modeli neuronowych dla potrzeb modelowania informacji medycznych rozciągniętych w czasie np. określania postępów niektórych chorób degeneracyjnych (choroba Alzheimera, choroba Parkinsona). Może to pomóc w doborze odpowiedniej terapii i sposobów postępowania nefarmakologicznego.

W pracy [99] autorzy prezentują możliwość wykorzystania modeli neuronowych do predykcji rehospitalizacji pacjentów po odbytych kursach rehabilitacyjnych, związanych z przebytym wylewem. Przyczyny dla której podjęta została próba modelowania powyższego zjawiska mają naturę kliniczną i ekonomiczną. Zjawisko koniecznej ponownej hospitalizacji po przebytych stanach o charakterze ostrym, jest pochodną wielu czynników i nie do końca wyjaśnioną. Dodatkowo w przypadku tak specyficznych stanów jak wylew i związana z tym konieczność hospitalizacji oraz ewentualnej rehospitalizacji pacjentów – generowane koszty stanowią poważne obciążenie budżetów płatników za świadczenia medyczne. Otrzymane systemy neuronowe były porównywane z klasycznymi modelami statystycznymi opartymi o technikę regresji logistycznej. Dane opisujące 9584 pacjentów ze 167 szpitali. Dane wejściowe wykorzystane do budowy porównywanych modeli zawierały m.in. informacje o charakterze demograficznym - wiek, płeć, rasa, status małżeński oraz klinicznym - okres hospitalizacji po wylewie, wartość FIM (*ang. Functional Independence Measure*). Parametr FIM opisujący zdolność do wypełniania codziennych obowiązków podzielony jest na sześć podgrup. Dane pierwotne podzielono na trzy równe części i każdy z uzyskanych podzbiorów był wykorzystywany jako plik uczący, walidacyjny i testowy (3-krotne wzajemne sprawdzanie). Wyjście miało charakter binarny (0;1) – przewidziana lub nie konieczność ponownej hospitalizacji.

Jako podstawową miarę dopasowania poszczególnych modeli wykorzystano pole powierzchni pod krzywą ROC (AUROC). W trakcie trenowania sieci wykorzystano algorytm wstecznej propagacji błędu z modyfikacją Langevina. Najlepszą znaną architekturą okazała się sieć

z jedną warstwą ukrytą zbudowaną z czterech neuronów. Wartość AUROC dla modelu neuronowego (0.74) okazała się lepsza niż metody opartej o technikę regresji logistycznej (0.68), choć różnica nie była istotna statystycznie. Autorzy we wnioskach zwracają uwagę na mniejszą bezwzględną liczbę nietrafnych klasyfikacji modelu neuronowego i sugerują, że metoda hybrydowa łącząca w sobie cechy sieci neuronowej i systemu logistycznego mogłaby wnieść wyraźną poprawę zdolności generalizacji systemu.

Podjęto również próby zastosowania sztucznych sieci neuronowych jako podstawy systemów skriningowych, mających za zadanie klasyfikowanie pacjentów będących nosicielami wirusa HIV lub z obecną pełnoobjawową chorobą AIDS. Spośród 1949 rekordów opisujących pacjentów autorzy wyselekcjonowali 1171 spełniających warunki włączenia do badania. Każdy w pacjentów opisany został z użyciem 9 zmiennych niezależnych. Specyficzne kodowanie sprawia, że ostateczny rozmiar wektora wejściowego wynosił 15 zmiennych (niektóre z wartości są podawane sieci z wykorzystaniem dwóch komórek wejściowych – jednej dla wartości pozytywnej, natomiast drugiej specyficznej dla wartości negatywnej). Wśród informacji podawanych na wejściu znalazły się: dane demograficzne (rasa, płeć) oraz o charakterze medycznym (ilość wizyt lekarskich, ambulatoryjnych, ilość dni hospitalizacji, ilość dni hospitalizacji w stanie ostrym) oraz droga narażenia. Mimo, że modelowano zmienną o charakterze binarnym autorzy zaproponowali aby warstwa wyjściowa zbudowana była z dwóch komórek, według podobnej metodyki jak opisana powyżej. Zbiór został podzielony na jedną parę uczący-testowy (odpowiednio 1026 oraz 145 rekordów), co było jedyną metodą kontroli zdolności generalizacji sieci.

Otrzymane wyniki oscylowały wokół granicy 65% pozytywnej klasyfikacji (wartości PPV – *positive predicted value* oraz NPV – *negative predicted value*). Mimo, że dobrana przez autorów metodyka może budzić pewne wątpliwości (specyficzne kodowanie danych, niewykorzystana technika 10-cio krotnego wzajemnego sprawdzania, opieranie całości

wniosków jedynie na jednym teście generalizacji, testowanie jedynie jednej funkcji aktywacji oraz niewykorzystanie architektur o więcej niż jedna warstwie ukrytej) na podkreślenie zasługuje fakt możliwości zastosowania sztucznych sieci neuronowych w tak skomplikowanym zagadnieniu jakim jest modelowanie stanu zdrowia pacjentów, gdzie możliwa liczba czynników wpływających na fazę choroby jest trudna do zdefiniowania.

Bogatą literaturę oraz konkretne zastosowania praktyczne posiadają sztuczne sieci neuronowe wykorzystywane w urologii [14,45,104,134]. Ciekawą propozycją jest ogólnie dostępne internetowe narzędzie pracujące w oparciu o sztuczną sieć neuronową [65]. Jego możliwości obejmują predykcję ryzyka przerzutów, w tym zarówno do otaczających węzłów chłonnych jak i przerzutów odległych oraz ryzyka nawrotu choroby po zabiegu prostatektomii i modelowanie czasu przeżycia dla pacjentów z nowotworem prostaty.

Opracowanie systemu wczesnej detekcji nowotworu prostaty w oparciu o modele neuronowe było podstawowym celem autorów opracowania [36]. Podstawowymi danymi klinicznymi wykorzystywanymi w trakcie uczenia sieci były: wiek pacjenta, całkowita wartość pomiaru poziomu antygenu sterczowego (*ang. PSA - prostate-specific antigen*), procent wolnego (nie związanego) PSA, objętość prostaty, objętość strefy przejściowej stercza (*ang. TZ – transition zone*), proporcje poziomu PSA do objętości gruczołu (PSAD) oraz objętości strefy przejściowej (PSA-TZ), współczynnik zmian wartości PSA w czasie (PSAV) oraz wartość określana jako DRE (*ang. digital rectal examination*), a więc oceniany przez urologa wynik badania *per rectum*. Po analizie wrażliwości sieci pierwotnych dokonano ostatecznego wyboru zmiennych kluczowych (preprocessing danych). Zbiór wynikowy podzielono na dwie kategorie, zależne od wyniku całkowitego poziomu PSA. Dla pacjentów z wynikami o wartościach w przedziale 2,5-4 ng/ml składnikami wektora wejściowego wykorzystanymi z dalszej analizy wykluczono: wiek, całkowity PSA, TZ, PSAV i DRE, natomiast w przypadku pacjentów o wartościach PSA w przedziale 4-10



ng/ml po analizie zrezygnowano ze zmiennych kodujących wiek, wartości DRE oraz TZ. Wyniki modeli neuronowych porównywano z modelem regresyjnym oraz prostymi modelami opartymi na pojedynczych wartościach diagnostycznych. Do badania zakwalifikowano 1246 pacjentów, z czego u 974 stwierdzono poziom PSA powyżej 4ng/ml, natomiast u 272 mężczyzn poddanych badaniu wartość PSA mieściła się w granicach 2,5-4 ng/ml. Obydwa zbiory podzielono na trzy grupy – uczącą, walidacyjną oraz testową (test zdolności generalizacji) stosunku 2:1:1. Autorzy nie zawarli w pracy informacji na temat stosowanego algorytmu i innych parametrów uczenia sieci poza wzmianką, że wykorzystywanym typem sieci był perceptron wielowarstwowy, natomiast w trakcie analizy wrażliwości korzystano z algorytmu genetycznego, mającego na celu optymalizację architektury sieci oraz wybór parametrów kluczowych.

Do oceny i porównania skuteczności poszczególnych modeli wykorzystano wartość pola powierzchni pod krzywą AUC (*ang. AUROC – area under operating curve*) oraz odsetki odpowiedzi pozytywnych dla diagnozy dodatniej oraz negatywnej (*ang. PPV – positive predictive value, NPV – negative predictive value*).

Analiza i predykcja specyficznych dla nowotworu prostaty czynników były celem pracy Gamita i współautorów [31,45]. Wśród ocenianych parametrów znalazły się dwa, których wartość ma bezpośredni wpływ na decyzję terapeutyczną. Określenie ryzyka wystąpienia przerzutów do regionalnych węzłów chłonnych (*ang. LNS – lymph node spread*), jest podstawą do podjęcia decyzji o zabiegu chirurgicznym – limfadenektomii. Drugim ocenianym parametrem jest wartość CP określająca stadium przenikania zmian nowotworowych poza prostatę (*ang. CP – capsular penetration*). Jego wartość jest czynnikiem pozwalającym ocenić skuteczność podjętej ścieżki terapeutycznej. W przedstawionym doniesieniu ocenianą metodą leczenia była brachyterapia. W przypadku pierwszego z opisywanych problemów autorzy wykorzystali bazę pacjentów zbudowaną z opisów 4133 przypadków jako zbioru uczącego. Testowanie zdolności generalizacji przeprowadzono wykorzystując dwie

inne bazy pacjentów pochodzących z innych ośrodków (odpowiednio 330 i 227 pacjentów). Zadanie postawione sieci polegało na sklasyfikowaniu pacjentów do grup o wysokim i niskim ryzyku przerzutów do regionalnych węzłów chłonnych. Dane uczące stanowiły: klasyfikacja kliniczna oparta o system cTNM, poziom antygenu sterczowego (PSA) oraz wyniki badania próbki tkanki pochodzącej z biopsji (*Gleason score*). Wykorzystując sieć bazująca na algorytmie wstecznej propagacji błędu, autorom udało się znaleźć architekturę modelu pozwalającą na osiągnięcie 98% skuteczności. Druga część doświadczenia polegała na określeniu stadium przenikania zmian nowotworowych poza prostatę. Wykorzystując zbiór danych zbudowany z 409 przypadków (348 zakwalifikowano do zbioru uczącego, natomiast pozostałe 61 przypadków to zbiór testowy) zbudowano i określono skuteczność modelu neuronowego. Niestety autorzy nie przeprowadzili pełnej walidacji systemu. Wektor uczący zbudowany był z informacji dotyczących wieku, rasy, poziomu antygenu sterczowego (PSA), wyników badań próbki tkanki pochodzącej z biopsji (*Gleason score*), klasyfikacji klinicznej opartą o system cTNM poszczególnych pacjentów. Modelowaną wartością było stadium rozwoju nowotworu obejmujące możliwość przenikania zmian nowotworowych (pTNM). W 83% przypadków udało się prawidłowo sklasyfikować pacjentów z grupy kontrolnej, wykorzystując model neuronowy zbudowany z wykorzystaniem algorytmu wstecznej propagacji błędu. Mimo uwag, które można skierować do autorów pracy (przede wszystkim – brak pełnej walidacji systemów np. z zastosowaniem 10-cio krotnego wzajemnego sprawdzania) wydaje się, że zaprezentowane wyniki pozwalają pozytywnie ocenić zastosowane narzędzia neuronowe.

Potter i współautorzy [104] podjęli próbę zastosowania sieci neuronowych, jako narzędzi predykcji ryzyka nawrotu choroby po zabiegu radykalnej prostatektomii. W trakcie doświadczeń testowano trzy możliwości grupowania danych uczących w celu zbadania wpływu rodzaju zmiennej sterującej na wynik. Wśród informacji podawanych na wejściu, były: wskaźniki kliniczne - próbki tkanki pochodzącej z biopsji (*Gleason*

score), informacje dotyczące rozmiaru prostaty oraz skuteczności zabiegu chirurgicznego, stadium zaawansowania; dane demograficzne – wiek; dane morfologiczne. Testowano specyficzny, bardzo ciekawy typ sieci neuronowych, których algorytm budowy i uczenia oparty jest na zjawiskach znanych z genetyki, choć w publikacji nie ma dokładniejszych informacji dotyczących jego szczegółów. Wyjściowy zbiór danych zawierający opis 214 przypadków podzielono losowo na trzy pary testowo-uczący, z których każda wykorzystywana była dla innego zbioru danych. Niestety autorzy nie zdecydowali się na przeprowadzenie pełnej walidacji systemów neuronowych. Modelowaną zmienną był nawrót choroby określany na podstawie poziomu antygenu sterczowego PSA, który jest klasycznym – choć nie do końca specyficznym – markerem raka prostaty. W zależności od rodzaju danych uczących systemy neuronowe z ok. 80% skutecznością potrafiły określić ryzyko nawrotu choroby. Oprócz odsetka pozytywnych klasyfikacji wśród ocenianych parametrów znajdowały się czułość, specyficzność i pole powierzchni pod krzywą ROC (AUROC). Wszystkie wyniki uzyskane dzięki zastosowaniu systemów neuronowych były istotnie statystycznie lepsze niż standardowe techniki statystyczne (regresja Coxa i Kaplana-Meiera).

#### 2.3.2.2 Modelowanie wskaźników epidemiologicznych

Cacciafesta wraz ze współpracownikami zastosowali sztuczne sieci neuronowe do modelowania odsetka dwuletniej przeżywalności u ludzi w podeszłym wieku [22]. Bazując na danych demograficznych, klinicznych (choroby przewlekłe), wynikach badań biochemicznych oraz dotyczących farmakoterapii i z wykorzystaniem klasycznej sieci neuronowej uczonej z wykorzystaniem algorytmu wstecznej propagacji błędu (logistyczna funkcja aktywacji) osiągnięto 85% skuteczność w teście generalizacji. Ortiz [97] zaproponował zastosowanie sztucznych sieci neuronowych do przewidywania jednorocznej śmiertelności u pacjentów z niewydolnością serca. Na podstawie informacji klinicznych oraz danych echokardiograficznych modelowano szansę przeżycia jednego roku. Baza

danych zawierała opis 95 pacjentów, których podzielono na grupy badaną oraz kontrolną. Wartość wyjściowa miała charakter binarny. Rezultaty uzyskane z zastosowaniem sieci neuronowych porównywano z rezultatami uzyskanymi na podstawie reguł zbudowanych przez klinicystów oraz przy zastosowaniu liniowej analizy dyskryminacyjnej (*ang. Linear Discriminant Analysis - LDA*). Najlepsza znaleziona architektura sieci wykazywała 90% skuteczność (71,4% czułość/93% specyficzność), co było wynikiem istotnie statystycznie lepszym w porównaniu do reguł klinicznych (82% skuteczność - 46,7% czułość/88% specyficzność) oraz metody LDA (67,4% skuteczność/67,5% czułość/27,8% specyficzność). W publikacji [37] Drożdż i współautorzy zaproponowali wykorzystanie sieci neuronowych do przewidywania współczynnika umieralności u pacjentów z ostrą niewydolnością wieńcową. Bazując na danych opisujących pacjentów pod kątem demograficznym, klinicznym (czynniki ryzyka, powikłania w trakcie hospitalizacji) oraz wdrożonej farmakoterapii testowano możliwość predykcji umieralności wśród hospitalizowanych pacjentów. Autorom udało się uzyskać 90% zgodność w teście generalizacji. Najprawdopodobniej ze względu na swój charakter (doniesienie konferencyjne) publikacja nie informuje na jakiej zasadzie spośród wektora pierwotnego wyselekcjonowano 13 zmiennych kluczowych. Zastrzeżenia może budzić również brak wdrożonej techniki 10-cio krotnej wzajemnej walidacji. Wśród innych zastosowań wymienić można modelowanie przeżywalności [119] oraz zachorowalności [113,120].

## 2.4 Regresja logistyczna

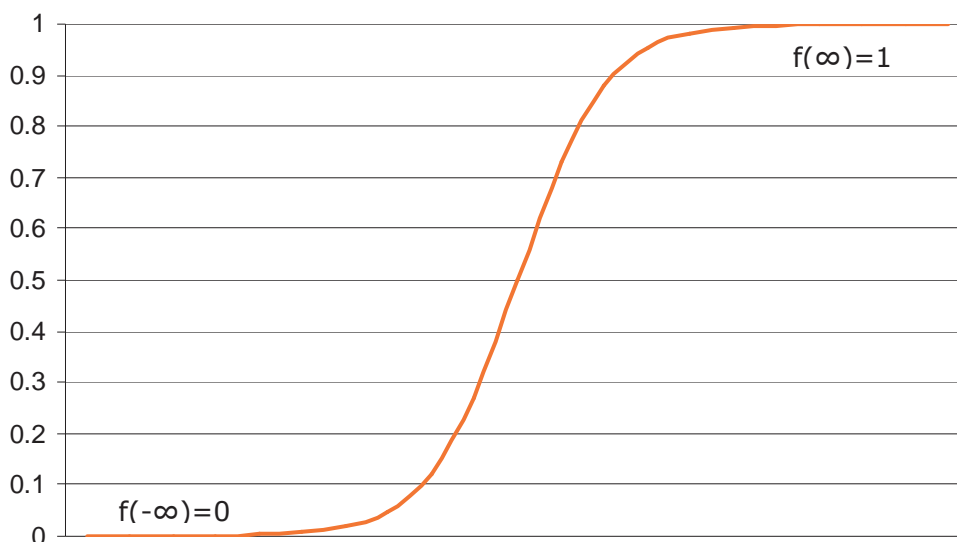
Regresja logistyczna jest jedną z metod modelowania nieliniowego. Powyższa technika jest wykorzystywana w przypadku gdy zmienna zależna ma charakter dychotomiczny [57]. Zależność między zmiennymi niezależnymi, a binarną zmienną zależną określa funkcja logistyczna, której postać przedstawia Równanie 22 [121].

**Równanie 22** Postać funkcji regresji logistycznej.

$$f(x) = \frac{e^x}{1 + e^x}$$

Schemat 20 przedstawia graficzną prezentację funkcji logistycznej.

**Schemat 20** Wykres funkcji logistycznej.



Budując model regresji logistycznej (wykorzystujący funkcję logistyczną) przyjmuje się założenie o dychotomicznym charakterze zmiennej wyjściowej. Wynikiem działania modelu jest prawdopodobieństwo wystąpienia jednej z dwóch możliwych odpowiedzi (0 lub 1).

Logistyczny model regresji przedstawia Równanie 23.

**Równanie 23** Zapis matematyczny logistycznego modelu regresji.

$$P(Y=1|x_1, x_2, \dots, x_n) = \frac{e^{(a_0 + \sum_{i=1}^k a_i x_i)}}{1 + e^{(a_0 + \sum_{i=1}^k a_i x_i)}}$$

gdzie:

$P(Y=1|x_1, x_2, \dots, x_n)$  – warunkowe prawdopodobieństwo przyjęcia wartości 1 dla zmiennych  $x_1, x_2, \dots, x_n$ .

$a_0, a_1, \dots, a_n$  – współczynniki regresji

Poszukiwanie postaci funkcyjnej modelu polega na oszacowaniu współczynników regresji, bazując na dostępnych danych. Najczęściej stosowaną metodą estymacji współczynników jest metoda największej wiarygodności (*ang. maximum likelihood*). Polega ona na maksymalizacji funkcji wiarygodności lub jej logarytmu.

Wykorzystując modele regresyjne jako narzędzia porównawcze dla modeli neuronowych zawsze korzystano ze wszystkich zmiennych, bez względu na istotność statystyczna poszczególnych współczynników równania funkcji. Oznacza to, że wszystkie zmienne sterujące były włączane w trakcie testowania. Powyższa procedura została przyjęta ze względu na konieczność zachowania porównywalnych warunków dla modeli neuronowych i statystycznych.

### 3 Cel pracy

Celem pracy było potwierdzenie możliwości zastosowanie sztucznych sieci neuronowych jako uniwersalnych narzędzi modelowania w epidemiologii i farmakoeconomic. Z całego spektrum nauk medycznych wymienione powyżej dyscypliny, są szczególnie predestynowane do poszukiwania nowych metod modelowania.

W celu udowodnienia powyższej tezy konieczne jest zarówno zbudowanie odpowiednich modeli neuronowych jak i ich przetestowanie oraz porównanie z modelami otrzymanymi z wykorzystaniem innych – odnośnikowych - metod modelowania. Ze względu na charakter wykorzystywanych danych wyniki modelowania neuronowego porównywane były z wynikami uzyskiwanymi metodą regresji logistycznej.

Ze względu na odmienną oczekiwanego wyniku przydatność sieci neuronowych ilustrowana jest osobno dla poszczególnych dyscyplin z wykorzystaniem odmiennych zbiorów danych. Budowa wykorzystywanych baz danych oraz zakres informacji zawartych w poszczególnych zbiorach determinowały ich zastosowanie.

Jako końcowy efekt prowadzonych prac zaplanowano gotowe narzędzia umożliwiające zastosowanie projektowanych systemów w codziennej praktyce farmaceutycznej i medycznej.

## 4 Oprogramowanie i sprzęt obliczeniowy

Aby móc symulować rozproszone przetwarzanie informacji na komputerze klasy PC niezbędne jest specjalistyczne oprogramowanie.

### 4.1 Oprogramowanie

Podstawowym narzędziem jest system *Nets2004* [86]. Jest to system do symulacji, testowania i analizy sieci neuronowych pracujący pod kontrolą środowisk Windows (9x, NT, 2000, XP). Graficzny, intuicyjny interfejs użytkownika pozwala na szybkie projektowanie zarówno architektury jak i parametrów sieci neuronowej. Symulator został napisany w języku Delphi 5.0 firmy Borland w technice obiektowej. Optymalizacja kodu doprowadziła do zmniejszenia wymagań wobec pamięci operacyjnej do tego stopnia, że *Nets2004* pozwala na pracę ze zbiorami danych o ilości rekordów dochodzących nawet do 5 000 000 (w zależności od rozmiaru wektora wejściowego i wyjściowego), na komputerze wyposażonym w 512 MB pamięci RAM. Takich możliwości nie posiada wiele komercyjnych programów w tej dziedzinie. Wielostopniowa automatyzacja pozwala na pracę systemu bez nadzoru operatora przez praktycznie dowolnie długi czas, konieczne jest jedynie stworzenie odpowiednie plików konfiguracyjnych w dowolnym edytorze tekstu. Aktualnie dostępne są następujące algorytmy uczenia sieci:

- propagacja wsteczna i jej modyfikacje (momentum, delta-bar-delta) [39,55,123]
- kontrapropagacja (algorytm WTM oraz gaz neuronowy) [123]
- lokalne przestrajanie [86]
- modyfikacje metody defuzyfikacji w uczeniu sieci rozmytych propagacją wsteczną (COA, MOM, M-SLIDE) [39]
- modyfikacja propagacji wstecznej dla sieci częściowo rekurencyjnych budowanych za pomocą tzw. warstw kontekstu (sieci typu Jordana-Elmana) [86]



W badaniach oprócz systemu *Nets2004* stosowanych jest kilka specjalistycznych programów narzędziowych [86]:

- *Learn\_conv, Split\_file* - systemy automatycznego tworzenia par zbiorów uczących i testowych z pojedynczych plików tekstowych, stosowany w metodzie walidacji krzyżowej (ang. cross-validation)
- *Eddan3* - arkusz kalkulacyjny do wstępnej obróbki danych na potrzeby sieci
- *Classif\_comp* - zaawansowany system analizy wyników otrzymanych podczas testowania modeli neuronowych

W trakcie analizy danych epidemiologicznych i prób ich wykorzystania do tworzenia modeli neuronowych konieczne okazało się stworzenia specjalistycznego narzędzia umożliwiającego konwertowanie informacji na postać zrozumiałą dla symulatora *Nets2004*. *Phmc\_app* to zaawansowany system umożliwiający także ingerencję w sposób kodowania danych, ich klasyfikację i podział informacji [101]. W trakcie tworzenia modeli referencyjnych (testowych) opartych o techniki regresyjne oraz analizy części wyników wykorzystano oprogramowanie *SigmaPlot*<sup>®</sup> 8.0 firmy SPSS Inc. oraz oprogramowanie MS Excel 2000 – składnik pakietu MS Office 2000 Professional.

## 4.2 Sprzęt obliczeniowy

Wszystkie obliczenia oraz analizy zostały wykonane na dwuprocessorowych stacjach roboczych, których konfiguracja przedstawiona jest w Tabeli 3.

**Tabela 3** Konfiguracja stacji roboczych wykorzystywanych w trakcie prac.

Stacja robocza	CPU	Pamięć RAM	HDD	System operacyjny
1	2x Intel Xeon 3,4 GHz	2048 Mb	36 Gb	Mandrake Linux
2	2x Intel Xeon 3,4 GHz	1024 Mb	36 Gb	Mandrake Linux
3	2x Intel Xeon 2,8 GHz	1024 Mb	36 Gb	Mandrake Linux

## 5 Zastosowanie sztucznych sieci neuronowych w farmakoekonomice

Jednym z wyzwań stojących przed farmakoekonomiką jest możliwość wykorzystania modeli przy określaniu efektu działania środka terapeutycznego. Pozwoliłoby to uniknąć części kosztownych procedur związanych z koniecznością przeprowadzania niektórych badań klinicznych. O ile przeprowadzenie odnośnych badań jest warunkiem *sine qua non* w przypadku wprowadzania nowego leku na rynek o tyle posiadając odpowiednia ilość zwalidowanych danych można pokusić się o zbudowanie modelu pozwalającego testować – w pewnym przedziale ufności – zakres dawek i wpływ ich zmian na efekt terapeutyczny oraz równoległe kombinacje wielolekowe. Powyższa procedura mogłaby się okazać szczególnie przydatna w przypadku schorzeń wymagających stosowania zamierzonej politerapii np. schorzeń nowotworowych. Otrzymane wyniki byłyby surogatem dla analiz farmakoekonomicznych. Po dokonaniu analizy kosztów zgodnie z wymogami przedstawionymi w rozdziale 2.2.4.1 możliwe jest przeprowadzenie pełnej analizy farmakoekonomicznej w oparciu o zamodelowane informacje. Jakość tak zaprojektowanej analizy i wyciąganych na jej podstawie wniosków zależy w największej mierze od jakości modelu neuronowego wyrażonej w postaci zdolności predykcyjnych.

### 5.1 Metodyka

#### 5.1.1 Dane

Informacje służące do budowy przykładowego systemu predykcyjnego oraz jego zastosowania w analizie farmakoekonomicznej są własnością Pracowni Farmakoepidemiologii i Farmakoekonomiki. Zostały wykorzystane do przeprowadzenia analizy farmakoekonomicznej schematów chemoterapii niedrobnokomórkowego raka płuc [120]. Dane pochodziły z czterech szpitali specjalistycznych, w których prowadzi się

terapię nowotworów płuc. Szczegółowy wykaz szpitali oraz liczbę pacjentów przedstawia Tabela 4.

**Tabela 4** Wykaz szpitali, z których pochodziły dane wykorzystane w trakcie badań.

Miasto	Szpital	Liczba pacjentów pierwotna	Liczba pacjentów ostateczna
Gdańsk	Klinika Onkologii i Radioterapii Akademii Medycznej	17	16
Poznań	Wielkopolskie Centrum Leczenia Gruźlicy i Chorób Płuc (Oddział Chemioterapii)	15	15
Warszawa	Klinika Chorób Płuc Centrum Onkologii – Instytutu im. Marii Skłodowskiej – Curie	51	51
Wrocław	Dolnośląskie Centrum Leczenia Chorób Płuc i Gruźlicy (Oddział VIA i VIB)	18	18

Po analizie i weryfikacji informacji wykluczono jeden z rekordów, ze względu na niekompletny zakres informacji. Ostatecznie do fazy poszukiwania modelu zakwalifikowano dane 100 pacjentów. Każdy z pacjentów opisany jest przez 30 zmiennych. Dostępne informacje podzielić można na 3 podstawowe grupy:

- demograficzne
- dotyczące stosowanej terapii lekowej
- informujące o zabiegach diagnostycznych i terapeutycznych oraz pobycie w szpitalu

Powyższe zmienne przedstawione zbiorczo zawierają odpowiednio Tabela 5 oraz Tabela 6.

**Tabela 5.** Zmienne opisujące rekordy pacjentów włączonych do badania (niedrobnokomórkowy rak płuc, 30 wejść).

L.p.	Grupa	Zmienna	Opis
1.	demograficzne	Płeć	Płeć pacjenta
2.		Wiek	Wiek pacjenta
3.	farmakoterapia	Winorelbina	Całkowita dawka winorelbiny [mg]
4.		Etopozyd	Całkowita dawka etopozydu [mg]
5.		Gemcytabina	Całkowita dawka gemcytabiny [mg]
6.		Cis-platyna	Całkowita dawka cis-platyny [mg]
7.		Carboplatyna	Całkowita dawka karboplatyny [mg]
8.		Zofran	Całkowita dawka ondansetronu [mg]
9.		Dexametazon	Całkowita dawka deksametazonu [mg]
10.		Navoban	Całkowita dawka tropisetronu [mg]
11.		Neupogen 300	Całkowita dawka filgrastimu 300 [mg]
12.		Neupogen 480	Całkowita dawka filgrastimu 480 [mg]
13.		Metoclopramid	Całkowita dawka metoklopramidu [mg]
14.	diagnostyka + hospitalizacja	Wizyty ambulatoryjne	Całkowita liczba wizyt ambulatoryjnych
15.		Dni w szpitalu	Całkowita liczba dni spędzonych w szpitalu
16.		RTG klatki piersiowej	Całkowita ilość badań RTG klatki piersiowej
17.		RTG kości	Całkowita ilość badań RTG kości
18.		Morfologia	Całkowita ilość wykonanych morfologii krwi
19.		Biochemia	Całkowita ilość badań biochemicznych krwi
20.		Krzepnięcie	Całkowita ilość badań krzepnięcia krwi
21.		Scyntygrafia	Całkowita ilość badań scyntygraficznych
22.		EKG	Całkowita ilość badań EKG
23.		TK klatki piersiowej	Całkowita ilość badań tomograficznych klatki
24.		TK głowy	Całkowita ilość badań tomograficznych głowy
25.		USG jamy brzusznej	Całkowita ilość badań USG jamy brzusznej
26.		Mocz	Całkowita ilość badań moczu
27.		NMR głowy	Całkowita ilość badań NMR głowy
28.		Ilość nawodnień	Całkowita ilość nawodnień w przypadku cis-platyny
29.		Ilość dni w OTZ	Całkowita ilość dni w OTZ
30.		Ilość frakcji promieniowania	Całkowita ilość otrzymanych frakcji promieniowania

**Tabela 6** Opis statystyczny zmiennych wejściowych.

L.p.	Zmienna	Ilość pacjentów objętych procedurą	Wartość minimalna	Wartość maksymalna	Wartość średnia	Odchylenie standardowe
1.	Płeć	81 M, 19 K				
2.	Wiek	100	41	77	62.1	8.98
3.	Winorelbina [mg]	34	100	898	305.9	172.55
4.	Etopozyd [mg]	26	300	4200	2255.9	1012.84
5.	Gemcytabina [mg]	90	10090	38400	27623.5	16522.97
6.	Cis-platyna [mg]	10	140	1380	635.4	315.88
7.	Carboplatyna [mg]	78	1500	3500	2435.0	666.69
8.	Zofran [mg]	74	8	368	59.6	60.06
9.	Dexametazon [mg]	34	4	252	59.9	53.29
10.	Navoban [mg]	5	5.0	150	59.1	45.03
11.	Neupogen 300 [mg]	3	300	2400	780.0	914.88
12.	Neupogen 480 [mg]	24	480	960	640.0	277.13
13.	Metoclopramid [mg]	72	20	335	95.2	85.78
14.	Wizyty ambulatoryjne	85	2	27	12.1	5.48
15.	Dni w szpitalu	87	2	129	40.3	28.64
16.	RTG	15	1	13	3.9	2.48
17.	RTG kości	100	1	5	1.4	1.12
18.	Morfologia	100	3	35	13.1	7.72
19.	Biochemia	44	1	19	6.4	3.00
20.	Krzepnięcie	17	1	12	3.3	2.94
21.	Scyntygrafia	38	1	2	1.2	0.39
22.	EKG	45	1	10	3.4	2.20
23.	TK klatki piersiowej	17	1	5	1.6	1.00
24.	TK głowy	54	1	2	1.4	0.49
25.	USG jamy brzusznej	49	1	8	2.3	1.94
26.	Mocz	9	1	7	3.3	2.14
27.	NMR głowy	90		2	1.1	0.33
28.	Ilość nawodnień	20	1	18	5.2	2.85
29.	Ilość dni w OTZ	33	1	52	17.8	16.16
30.	Ilość frakcji promieniowania	40	1	40	21.4	12.60

### 5.1.2 Wartość wyjściowa (modelowana)

Informacją wyjściową, podlegającą modelowaniu była wartość liczbową opisująca długość przeżycia pacjenta od momentu zdiagnozowania choroby wyrażona w tygodniach. Informacje opisujące wartość wyjścia opisane są w Tabeli 7.

**Tabela 7** Opis statystyczny wartości modelowanej (przeżycia).

Wartość modelowana	Wartość minimalna [tyg.]	Wartość maksymalna [tyg.]	Wartość średnia [tyg.]	Odchylenie standardowe [tyg.]
Przeżycie	7.57	120.71	35.87	22.37

### 5.1.3 Ocena błędu systemów neuronowych

Modelowana wartość wyjściowa miała charakter dyskretny (binarny lub przedziałowy) oraz ciągły. W zależności od jej właściwości stosowano różne schematy oceny procesu uczenia (błędu popełnianego przez nauczoną sieć).

Kontrolę procesu uczenia sieci w przypadku zmiennej o charakterze ciągłym przeprowadzano stosując wartość RMSE (*ang. Root Mean Squared Error*).

**Równanie 24** Wzór umożliwiający obliczenie błędu średniokwadratowego (RMSE).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_p - y_o)^2}$$

gdzie:

$n$  – ilość rekordów zbioru testowego

$y_p$  – wartość podawana przez sieć na wyjściu

$y_o$  – wartość oczekiwana

W przypadku zmiennej o charakterze dyskretnym oceniano odsetek odpowiedzi prawidłowych (całkowity oraz cząstkowy dla poszczególnych klas). Standardowo próg pobudzenia, dla którego przeprowadzano analizę wynosił 0,5 (50% pobudzenia). Szukano również wartości optymalnej wartości progów definiowanej jako wartość dla której odsetek wszystkich odpowiedzi prawidłowych był najwyższy. Dodatkowo wykorzystywano wskaźnik oceny dopasowania modelu jakim jest wartość AUROC (*ang.*

*Area Under Receiver Operating Curve* lub *Area Under Receiver Operating Characteristic*) [17].

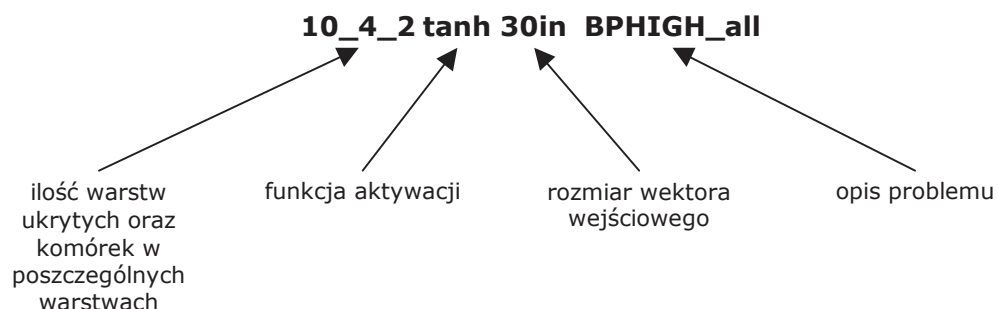
#### 5.1.4 Walidacja systemów neuronowych

Jedną z technik oceny generalizacji sieci i równocześnie eliminacji minimów lokalnych wykorzystywanych w opracowaniach jest tzw. k-krotna krzyżowa walidacja (*ang. k-fold cross-validation*) czyli k-krotne wzajemne sprawdzanie systemu. Jest to metoda polegająca na podziale zbioru danych na k par: uczący-testowy. Każda architektura sieci uczona jest następnie k razy - za każdym razem na innej parze zbiorów uczący-testowy. Standardowo zalecana wielkość k wynosi 10 (dziesięciokrotne wzajemne sprawdzanie), co pozwala ujednoczyć otrzymywane w różnych ośrodkach wyniki i ułatwia ich obiektywne porównanie.

#### 5.1.5 Schemat opisu testowanych architektur sieci neuronowych

W trakcie doświadczeń testowano setki architektur sieci, w związku z czym konieczne było stworzenie systemu opisu pozwalającego identyfikować architekturę sieci, funkcję aktywacji, rozmiar wektora wejściowego i badany problem. Stosowany sposób opisu przedstawiony został na Schemacie 21.

**Schemat 21** Sposób opisu testowanych architektur sieci neuronowych.



#### 5.1.6 Formatowanie (preprocessing) danych

Często w trakcie doświadczeń i budowania modeli opartych o sztuczne sieci neuronowe, podobnie jak dla innych empirycznych metod obliczeniowych konieczne jest wstępne przetwarzanie danych. Dotyczy to zarówno danych uczących, a więc prezentowanych sieci w trakcie budowy

modelu jak i danych wyjściowych uzyskiwanych w wyniku pracy sieci. Na opisywanym etapie badań wdrożono dwa sposoby wstępnego przetwarzania danych (Schemat 22). Jednym z nich była zmiana charakteru wartości wyjściowej (zmienna modelowana została przetworzona do wartości dyskretnej, odpowiednio binarnej i przedziałowej). Kolejny krok to zaszumienie danych pierwotnych.

#### 5.1.6.1. Zaszumianie danych

Zaszumianie pierwotnych danych uczących jest jednym ze sposobów wstępnego przetwarzania posiadanych informacji, podawanych sieci na wejściu. Podobna technika ma na celu zarówno zwiększenie ilości rekordów uczących (szczególnie w przypadku stosunkowo niewielkiego zbioru danych) jak i pełniejsze oddanie warunków rzeczywistych jeśli reprezentatywność zbioru uczącego jest niska. Końcowym efektem i celem podobnego postępowania jest wzrost skuteczności modelu neuronowego.

W trakcie doświadczeń przetestowano pięć schematów zaszumiania danych (Tabela 8), gdzie opcja multiplikacja oznacza ile razy dany rekord oryginalny był powielany (w sposób losowy), po zmianie o odpowiedni współczynnik wartości szumu.

**Tabela 8** Zakres zaszumiania danych pierwotnych.

Nr	Opcja zaszumiania	Multiplikacja	Wartość szumu
1.	szum, randomizacja	2	2%
2.	szum, randomizacja	2	5%
3.	szum, randomizacja	2	10%
4.	szum, randomizacja	5	5%
5.	szum, randomizacja	10	5%

#### 5.1.6.2. Dyskretyzacja

Dyskretyzacja polega na zastąpieniu danych ciągłych nowymi wartościami o charakterze dyskretnym, które odpowiadają przedziałom wartości ciągłych [29]. W wyniku procesu dyskretyzacji otrzymywana jest skończona liczba przedziałów, będąca najczęściej kompromisem między dokładnością sposobu reprezentowania informacji, a złożonością procesu



uczenia i uzyskiwanych hipotez. Dyskretyzacja jest często stosowanym w trakcie analizy informacji przetworzeniem przestrzeni danych, a najczęstszą przyczyną są trudności modelu z interpretacją zmiennych ciągłych. Prawidłowo przeprowadzona dyskretyzacja może skutkować poprawą efektywności modelu.

Podstawowy podział metod dyskretyzacji podaje dwie możliwe drogi jej wykonania – prostą oraz zaawansowaną. Metody zaliczane do prostych nie uwzględniają rozkładu atrybutów w poszczególnych klasach, natomiast metody zaawansowane reagują na oczekiwania związane z zakresem informacji zawartych w zbiorach uczących.

Najprostszym schematem dyskretyzacji danych ciągłych jest binaryzacja, a więc przefiltrowanie posiadanych informacji przez pewną wartość progową i podział na dwa przedziały. W trakcie prac testowano dwie progowe wartości wyjścia (czasu przeżycia) – 30 oraz 35 tygodni. Oznacza to, że wszystkie wartości oznaczające czas przeżycia poniżej odpowiednio 30 i 35 tygodni zostały przekodowane do wartości 0, natomiast powyżej progów do wartości 1. Testowane wartości progowe zostały dobrane według następujących algorytmów:

- przyjęcie progu 30-to tygodniowego umożliwiła otrzymanie równej liczby przypadków dla obu grup (poniżej i powyżej wartości progowej)
- 35 tygodni to średni literaturowy czas przeżycia pacjentów w IIIB i IV stadium zaawansowania niedrobnokomórkowego raka płuc (mediana przeżycia) [73,84,100,125,126,127]

**Tabela 9** Liczebność poszczególnych klas przy podziale binarnym z zastosowaniem progu 30 i 35 tygodni przeżycia.

Próg [tyg.]	Liczba 0	Liczba 1
30	52	48
35	63	37

W trakcie doświadczeń wykorzystano również inny sposób dyskretyzacji, który mógł okazać się szczególnie przydatny ze względu na

stosunkowo niewielką ilość danych. Każdy rekord wszystkich zbiorów uczących oraz odpowiadających im zbiorów testowych został rozwinięty do 115 wierszy. Powyższa liczba powstała poprzez porównanie wartości maksymalnej (121-no tygodniowe) oraz minimalnej (7-mio tygodniowe przeżycie) na wyjściu dla danych pierwotnych. Następnie dla wartości kodującej przeżycie poniżej nominalnej dla danego rekordu przypisano wartość 1, natomiast dla wartości wyższej – 0. Ostatecznie każdy ze zbiorów uczących zawierał 10 350 wierszy, natomiast testowy – 1150. Dodatkową konsekwencją była zmiana ilości zmiennych wejściowych z 30 do 31. Metodyka podziału na pary testowy-uczący była odmienna od podejścia standardowego. W pierwszym etapie z danych pierwotnych (100 rekordów) uzyskiwano dziesięć par (zgodnie z metodyką 10-cio krotnego wzajemnego sprawdzania) na drodze losowego przydziału do odpowiedniego zbioru, które następnie przetwarzano z wykorzystaniem makra do postaci o 31 wejściach. Mimo modelowania wyjścia o charakterze binarnym wynik pracy sieci (testowania generalizacji) oceniany był z wykorzystaniem wartości RMSE, ponieważ po etapie testowania dekodowano rekordy do wartości ciągłej, wykorzystując przygotowane w tym celu makro.

Przeprowadzono również testy z większą ilością zbiorów. Dokonano podziału na odpowiednio 3 oraz 4 klasy i testowano wpływ powyższej modyfikacji zmiennej wyjściowej na otrzymywany wynik. Tabela 10, Tabela 11, Tabela 12, Tabela 13 przedstawiają odpowiednio zakresy przedziałów przy podziale na cztery i trzy grupy oraz ich liczebności. W obydwu przypadkach testowano różne wartości długości przedziału aby skontrolować wpływ doboru parametrów dyskretyzacji (algorytmu podziału na klasy) na zdolność generalizacji modelu.

**Tabela 10** Liczebność poszczególnych klas przy podziale na cztery klasy (długość pierwszego przedziału – 15 tyg.).

Początek przedziału	Koniec przedziału	Numer przedziału	Liczba rekordów
7	22	1	27
22	37	2	39
37	52	3	15
52	121	4	19
SUMA			100

**Tabela 11** Liczebność poszczególnych klas przy podziale na cztery klasy (długość pierwszego przedziału – 35 tyg.).

Początek przedziału	Koniec przedziału	Numer przedziału	Liczba rekordów
7	42	1	71
42	77	2	23
77	112	3	5
112	121	4	1
SUMA			100

**Tabela 12** Liczebność poszczególnych klas przy podziale na trzy klasy (długość pierwszego przedziału – 15 tyg.).

Początek przedziału	Koniec przedziału	Numer przedziału	Liczba rekordów
7	22	1	27
22	37	2	39
37	121	3	34
SUMA			100

**Tabela 13** Liczebność poszczególnych klas przy podziale na trzy klasy (długość pierwszego przedziału – 35 tyg.).

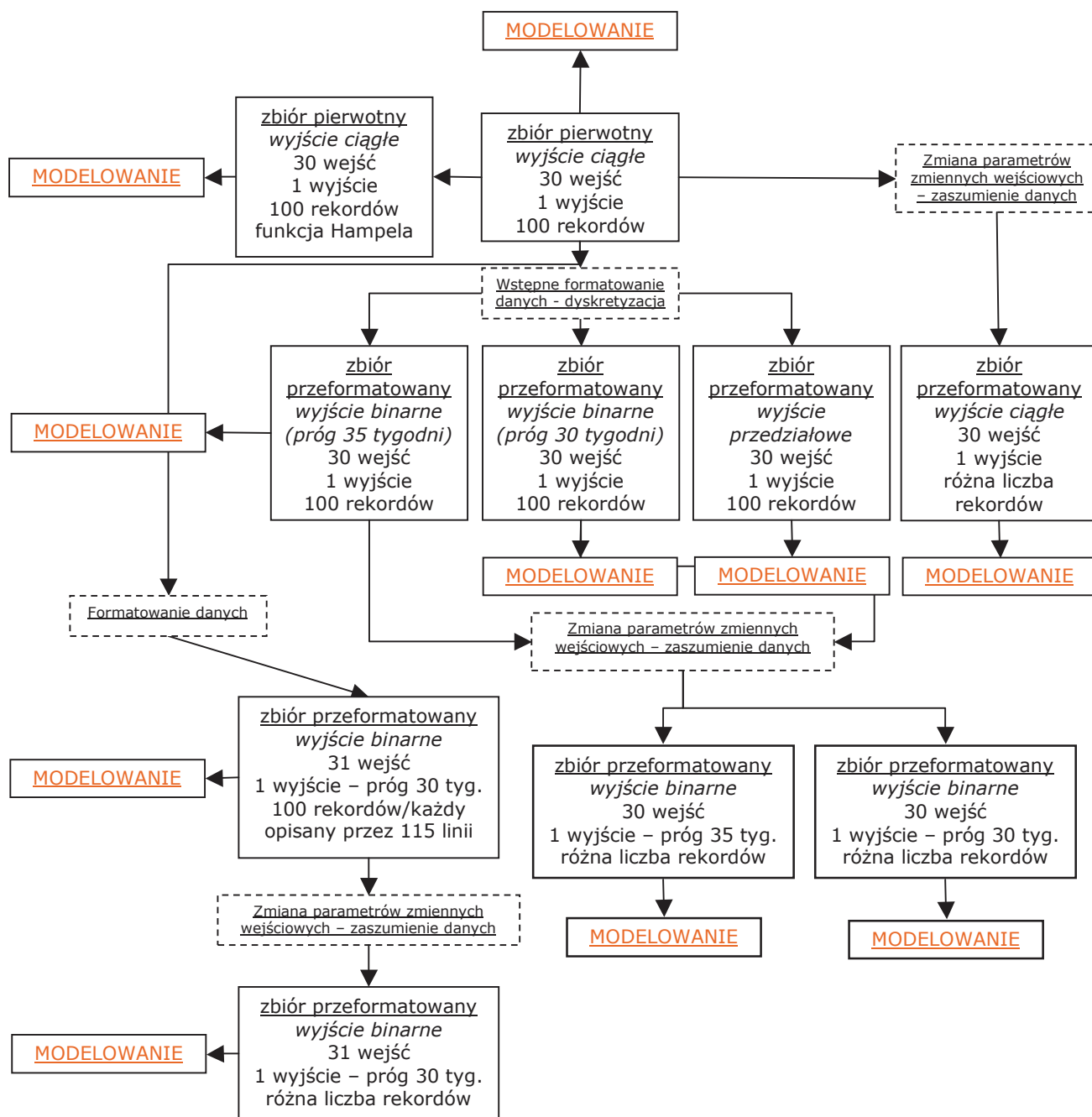
Początek przedziału	Koniec przedziału	Numer przedziału	Liczba rekordów
7	42	1	71
42	77	2	23
77	121	3	6
SUMA			100

15-to i 35-cio tygodniowe długości przedziałów zostały dobrane w ten sposób aby spełnić dwa kryteria: zbalansowanie liczebności klas (35 tygodni) oraz oddanie realnego odsetka populacji chorych w stadium IIIB oraz IV przeżywających określony przedział czasu.

#### 5.1.7 Przebieg pracy

Ze względu na specyfikę danych wykorzystywanych jako dane uczące dla modeli neuronowych oraz małą ilość rekordów, modelowanie na danych pierwotnych nie było jedyną wykorzystywaną drogą. Przeprowadzono doświadczenia, w których dane pierwotne poddano wstępnemu przetwarzaniu (*ang. preprocessing*) różnymi metodami przedstawionymi szczegółowo w rozdziale 5.1.6. Celem powyższych działań było przetestowanie różnego sposobu prezentacji danych uczących i ocena ich wpływu na wynik końcowy modelu (zdolność predykcji). Schemat 22 prezentuje plan kolejnych metod wstępnego przetwarzania danych, które dokładnie opisane są w kolejnych rozdziałach.

**Schemat 22** Zakres modyfikacji danych pierwotnych - diagram przebiegu doświadczeń.



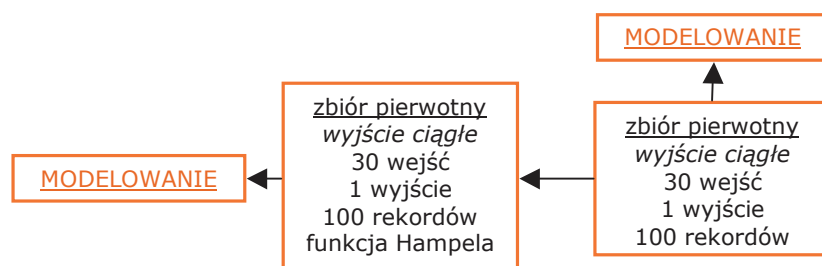
Wszelkie dodatkowe modyfikacje i specyficzne rozwiązania konieczne do wprowadzenia w danym etapie pracy opisywane są na bieżąco w poszczególnych rozdziałach. Poszczególne etapy pracy są wyróżnione na schematach czerwoną pogrubioną linią.

## 5.2 Rezultaty poszczególnych etapów pracy

### 5.2.1 Modelowanie z wykorzystaniem danych pierwotnych – wyjście ciągłe

Jak przedstawia Schemat 23 dane pierwotne zawierały 100 rekordów, z których każdy opisany był przez 30 zmiennych. W każdym przypadku modelowano jedną zmienną wyjściową, która miała charakter ciągły (Tabela 7).

**Schemat 23** Etap I - modelowanie z wykorzystaniem danych pierwotnych



W trakcie doświadczeń przetestowano architektury, które przedstawia Tabela 14.

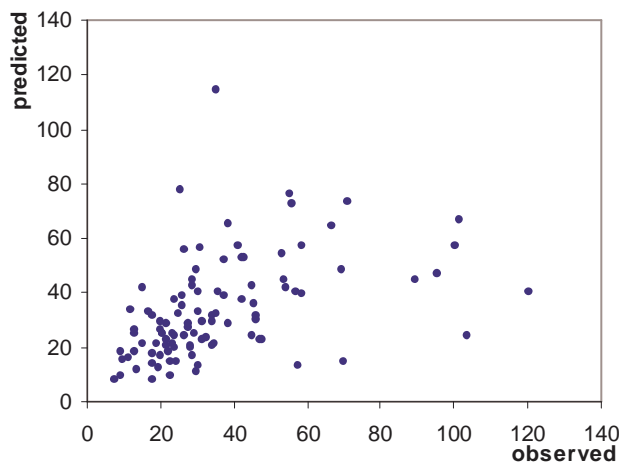
**Tabela 14** Architektury testowane w trakcie doświadczeń na danych pierwotnych – najlepsze uzyskane wyniki.

Nr	Architektura sieci	Wartość RMSE
1.	2hid M-SLIDE_30in_all	20.97
2.	10_7_5_2hid sigma_30in_all	21.36
3.	20_10_2hid sigma_30in_all	23.04
4.	10hid M-SLIDE_30in_all	24.29
5.	12_9_7_5_2hid sigma_30in_all	24.39
6.	3hid M-SLIDE_30in_all	24.42
7.	100hid sigma_30in_all	24.56
8.	2hid tanh_30in_all	25.59
9.	2hid sigma_30in_all	25.74
10.	7_5_3hid sigma_30in_all	26.59

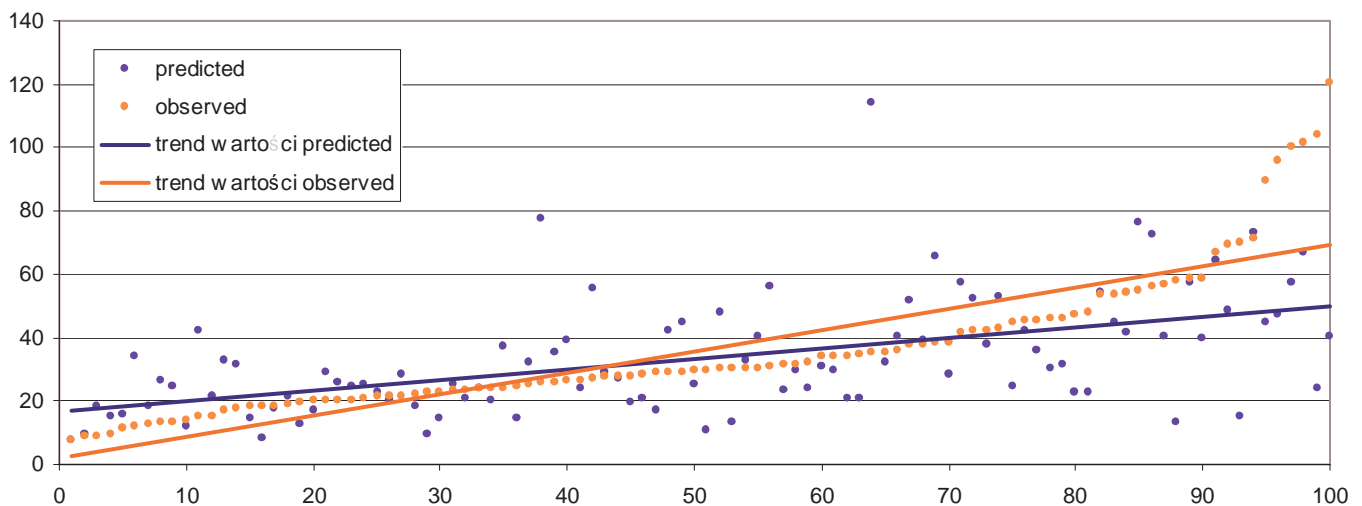
Najniższa uzyskana na tym etapie wartość RMSE wynosiła 20.97. Nie jest to wartość zadawalająca, która pozwalałaby potwierdzić hipotezę o prawidłowym odwzorowaniu opisywanej rzeczywistości przez uzyskany model. Niemniej jednak analizując trendy obydwu wartości zauważyć

można, że system potrafi przewidzieć tendencję zmiany modelowanej wartości. Daje to asumpt do przeprowadzenia kolejnych etapów doświadczeń, jednak konieczne jest poszukiwanie modeli będących w stanie w sposób pełniejszy odwzorować obserwowaną wartość.

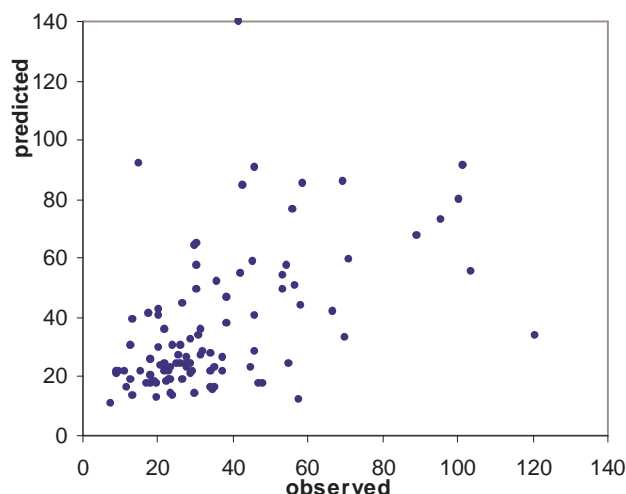
**Schemat 24** Wykres zależności wartości rzeczywistej do przewidzianej przez model dla architektury 2hid M-SLIDE, uczonej na danych pierwotnych z wyjściem ciągłym.



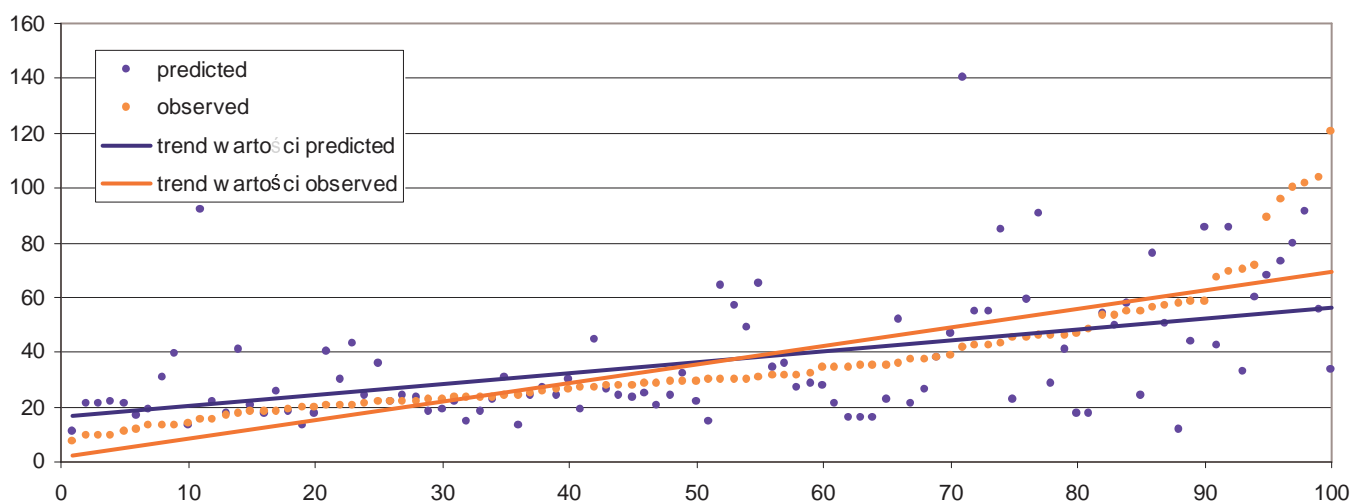
**Schemat 25** Porównanie wartości rzeczywistych i przewidzianych przez model dla architektury 2hid M-SLIDE, uczonej na danych pierwotnych z wyjściem ciągłym; analiza trendów.



**Schemat 26** Wykres zależności wartości rzeczywistej do przewidzianej przez model dla architektury 10\_7\_5\_2hid sigma, uczonej na danych pierwotnych z wyjściem ciągłym.



**Schemat 27** Porównanie wartości rzeczywistych i przewidzianych przez model dla architektury 10\_7\_5\_2hid sigma, uczonej na danych pierwotnych z wyjściem ciągłym; analiza trendów.



Doświadczenia polegające na zmianie charakteru funkcji błędu nie przyniosły spodziewanych efektów. Najlepsze rezultaty (najniższe wartości RMSE) uzyskane z wykorzystaniem funkcji Hampela oscylowały w granicach 30 (Tabela 15).



**Tabela 15** Architektury testowane w trakcie doświadczeń na danych pierwotnych z zastosowaniem funkcji Hampela – najlepsze uzyskane wyniki

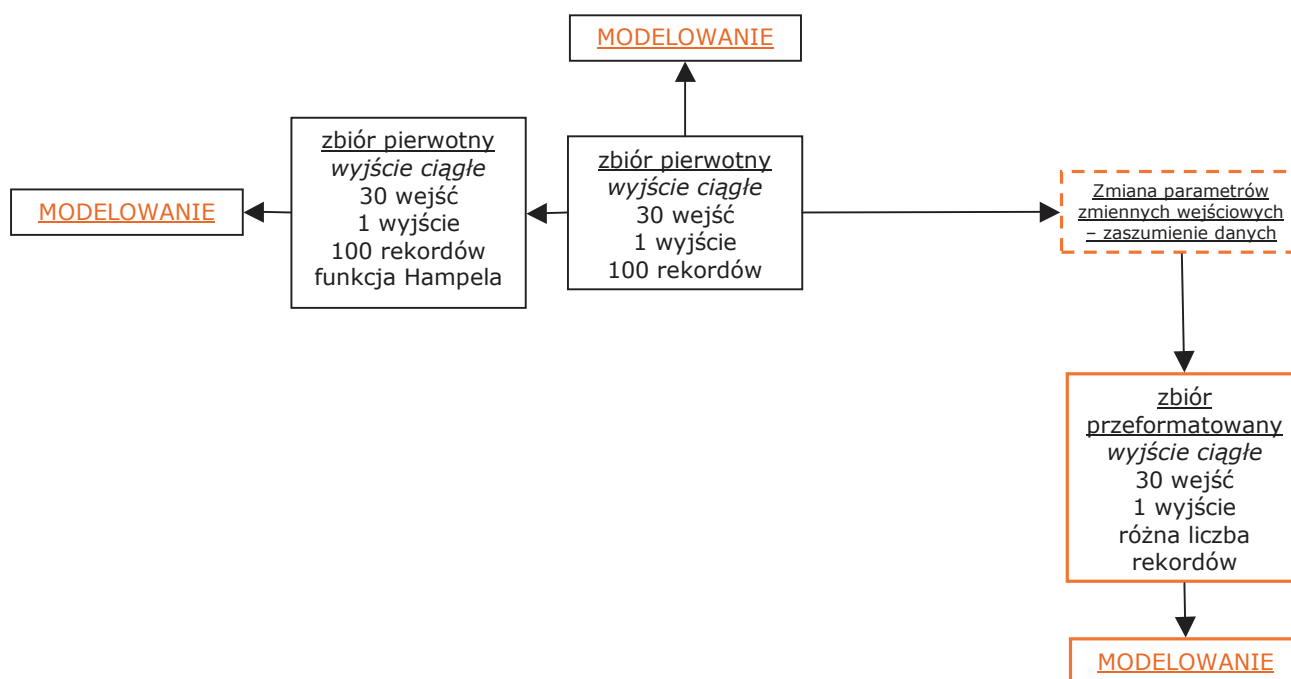
Nr	Architektura sieci	Wartość RMSE
1.	10hid sigma_30in_all_Hampel	29.09
2.	10_7hid sigma_30in_all_Hampel	29.76
3.	10hid tanh_30in_all_Hampel	30.71
4.	10hid fsr_30in_all_Hampel	31.45
5.	10_7hid tanh_30in_all_Hampel	32.26
6.	10_7hid fsr_30in_all_Hampel	33.63

## 5.2.2 Modelowanie z wykorzystaniem danych przeformatowanych

### 5.2.2.1 Dane zaszumione - wyjście ciągłe

Kolejny etap pracy polegał na przetestowaniu możliwości poprawy wyników sieci uczonych na danych z dodanym losowo szumem. Charakter wyjścia został zachowany (wyjście ciągłe - Schemat 28).

**Schemat 28** Etap II – wstępne formatowanie danych oraz modelowanie z wykorzystaniem danych przeformatowanych – zaszumienie danych.



Efektem dodania losowego szumu do danych uczących jest zwiększenie liczby rekordów wykorzystywanych w trakcie trenowania sieci.

**Tabela 16** Parametry zaszumiania danych; ilość rekordów zbiorów wynikowych – wyjście ciągłe.

Nr	Opcja zaszumiania	Multiplikacja	Wartość szumu	Ilość rekordów zbioru uczącego
1.	szum, randomizacja	2	2%	450
2.	szum, randomizacja	2	5%	450
3.	szum, randomizacja	2	10%	450
4.	szum, randomizacja	5	5%	990
5.	szum, randomizacja	10	5%	1890

W trakcie doświadczeń przetestowano 53 architektury sieci uczonych na zbiorach zaszumianych z różnymi parametrami. Tabela 17, Tabela 18, Tabela 19, Tabela 20, Tabela 21 przedstawiają najlepsze uzyskane wyniki.

**Tabela 17** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem ciągłym – najlepsze wyniki dla zbioru o parametrach parametrów szumu 2x2%.

Nr	Architektura sieci	Wartość RMSE
1.	2hid M-SLIDE_30in_all	20.89
2.	20_16_8_4hid sigma_30in_all	23.83
3.	7_5_3hid tanh_30in_all	24.19
4.	20_16_8_4hid fsr_30in_all	24.35
5.	60_20_5hid fsr_30in_all	25.05
6.	2hid sigma_30in_all	25.73
7.	60_20hid tanh_30in_all	25.86
8.	20_10_2hid sigma_30in_all	26.13
9.	60_20_5hid sigma_30in_all	26.50
10.	20_18_10hid fsr_30in_all	26.57

**Tabela 18** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem ciągłym – najlepsze wyniki dla zbioru o parametrach parametrów szumu 2x5%.

Nr	Architektura sieci	Wartość RMSE
1.	2hid M-SLIDE_30in_all	21.06
2.	20_16_8_4hid sigma_30in_all	22.01
3.	60_20_5hid tanh_30in_all	23.61
4.	12_9_7_5_2hid sigma_30in_all	23.98
5.	20_18_10hid tanh_30in_all	24.27
6.	3hid M-SLIDE_30in_all	24.41
7.	5hid M-SLIDE_30in_all	24.57
8.	4hid M-SLIDE_30in_all	24.74
9.	20_10_2hid sigma_30in_all	24.84
10.	10_7_5_2hid sigma_30in_all	24.97

**Tabela 19** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem ciągłym – najlepsze wyniki dla zbioru o parametrach parametrów szumu 2x10%.

Nr	Architektura sieci	Wartość RMSE
1.	2hid M-SLIDE_30in_all	21.77
2.	20_10_2hid sigma_30in_all	22.54
3.	12_9_7_5_2hid sigma_30in_all	22.56
4.	18_8hid tanh_30in_all	22.96
5.	5_3hid sigma_30in_all	22.97
6.	7_5_3hid sigma_30in_all	23.36
7.	2hid tanh_30in_all	23.60
8.	20_16_8_4hid sigma_30in_all	23.64
9.	60_20_5hid fsr_30in_all	23.80
10.	10hid fsr_30in_all	23.98

**Tabela 20** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem ciągłym – najlepsze wyniki dla zbioru o parametrach parametrów szumu 5x5%.

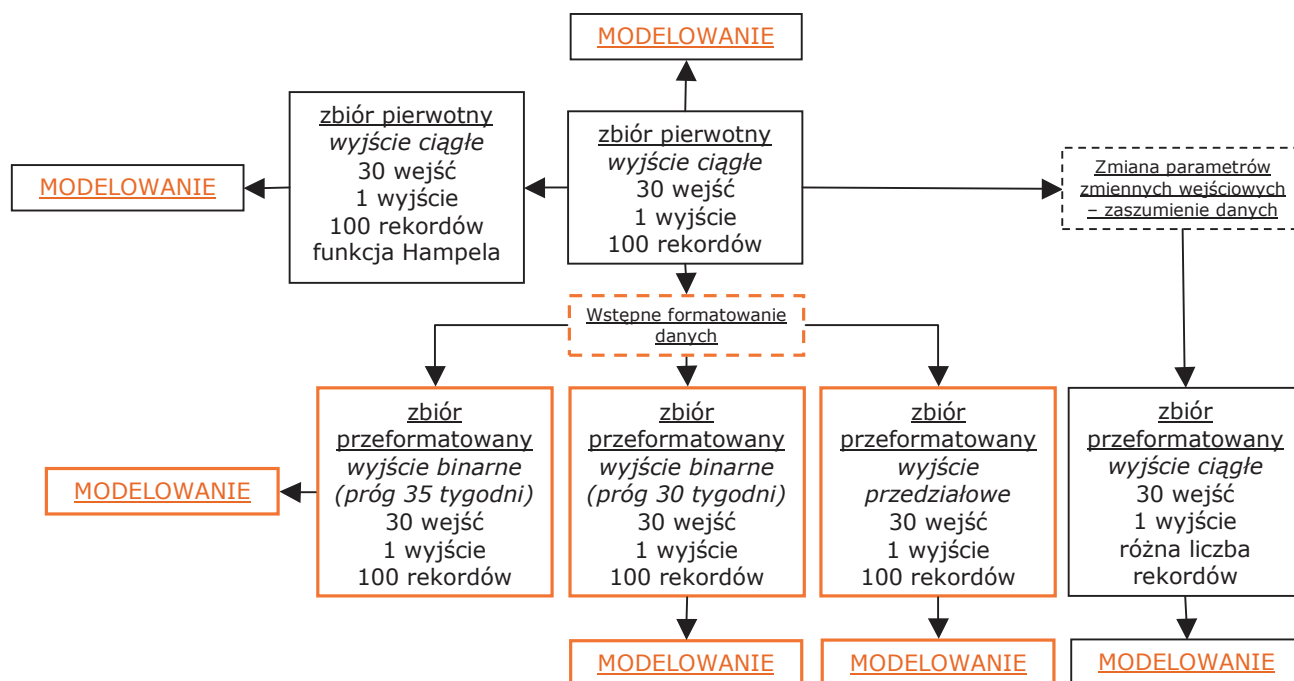
Nr	Architektura sieci	Wartość RMSE
1.	2hid M-SLIDE_30in_all	20.60
2.	12_9_7_5_2hid sigma_30in_all	20.77
3.	10_7_5_2hid sigma_30in_all	23.35
4.	20_10_2hid tanh_30in_all	23.38
5.	15_5hid tanh_30in_all	23.49
6.	20_10hid tanh_30in_all	23.76
7.	20_18_10hid tanh_30in_all	23.91
8.	7_5_3hid tanh_30in_all	24.14
9.	5_3hid sigma_30in_all	24.28
10.	7_5_3hid fsr_30in_all	24.34

**Tabela 21** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem ciągłym – najlepsze wyniki dla zbioru o parametrach parametrów szumu 10x5%.

Nr	Architektura sieci	Wartość RMSE
1.	2hid M-SLIDE_30in_all	21.04
2.	4hid M-SLIDE_30in_all	22.15
3.	20_10_2hid tanh_30in_all	22.52
4.	20_16_8_4hid sigma_30in_all	22.93
5.	20_18_10hid tanh_30in_all	23.05
6.	15_5hid tanh_30in_all	23.76
7.	12_9_7_5_2hid sigma_30in_all	24.04
8.	10hid M-SLIDE_30in_all	24.23
9.	7_5_3hid sigma_30in_all	24.41
10.	20_10_2hid sigma_30in_all	24.52

### 5.2.2.2 Zmienna wyjściowa dyskretna

**Schemat 29** Etap III – wstępne formatowanie danych oraz modelowanie z wykorzystaniem danych przeformatowanych – dyskretyzacja danych.



Jak przedstawiono powyżej, dyskretyzacja jest jednym ze sposobów rozwiązania problemu polegającego na niemożności dopasowania parametrów modelu do analizowanego zagadnienia.

Pierwszym etapem preprocessingu danych i zmiany sposobu prezentacji informacji sieci było zastosowanie dyskretyzacji do wartości przedziałowych. Testowano dwie podstawowe opcje – podział na trzy i cztery klasy. W obydwu przypadkach podziału dokonano w oparciu o dwa algorytmy wyznaczania granic przedziałów przyjmując granice przedziałów na 15 i 35 tygodni.

Tabela 22 przedstawia najlepsze uzyskane wyniki dla sieci trenowanych z wykorzystaniem danych z 3 klasami, przy domyślnej granicy klasy – 15 tygodni.

**Tabela 22** Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem przedziałowym (3 klasy; granica klasy – 15 tygodni) – najlepsze uzyskane wyniki.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		
			3	2	1
1.	15_5 hid tanh_30in_all	54	68	56	33
2.	5_3_2 hid sigma_30in_all	54	68	49	44
3.	60_20hid tanh_30in_all	53	68	49	41
4.	7_5_3hid tanh_30in_all	52	71	44	41
5.	10_7 hid tanh_30in_all	51	74	44	33
6.	20_18_10hid fsr_30in_all	51	59	46	48
7.	5_3 hid tanh_30in_all	50	71	36	44
8.	10_7 hid fsr_30in_all	49	53	62	26
9.	5_3_2 hid fsr_30in_all	49	53	62	26
10.	20_10_2 hid sigma_30in_all	48	65	33	48

Tabela 23 przedstawia najlepsze uzyskane wyniki dla sieci uczonych na danych z 3 klasami, przy domyślnej granicy klasy – 35 tygodni.

**Tabela 23** Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem przedziałowym (3 klasy; granica klasy – 35 tygodni) – najlepsze uzyskane wyniki.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		
			3	2	1
1.	5_3 hid tanh_30in_all	67	33	26	83
2.	10 hid sigma_30in_all	66	17	39	79
3.	7_5_3hid tanh_30in_all	65	17	39	77
4.	10_7 hid fsr_30in_all	63	17	35	76
5.	18_8 hid tanh_30in_all	63	17	26	79
6.	10_7 hid tanh_30in_all	63	17	22	80
7.	7_5_3hid sigma_30in_all	62	33	13	80
8.	20_10_2 hid sigma_30in_all	62	33	4	83
9.	20_10 hid fsr_30in_all	62	17	30	76
10.	5 hid sigma_30in_all	62	17	30	76

Ze względu na niezadowalające wyniki modelowania z wykorzystaniem danych przygotowanych w oparciu o schemat podziału na trzy klasy konieczne było przetestowanie możliwości uzyskania wiarygodnych modeli z wykorzystaniem innego algorytmu podziału na klasy. Kolejny etap obejmował modelowanie z wykorzystaniem zbioru z czterema klasami wyjścia. Podobnie jak wcześniej wykorzystano dwie wartości określające granica klasy – 15 i 35 tygodni.

**Tabela 24** Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem przedziałowym (4 klasy; granica klasy – 15 tygodni) – najlepsze uzyskane wyniki.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas			
			4	3	2	1
1.	20_18_10hid sigma_30in_all	38	58	27	26	48
2.	5_3 hid tanh_30in_all	38	58	33	36	30
3.	10_7 hid tanh_30in_all	37	47	33	36	33
4.	2 hid fsr_30in_all	36	11	33	49	37
5.	5_3 hid sigma_30in_all	35	47	20	23	52
6.	60_20hid fsr_30in_all	35	26	33	44	30
7.	18_8 hid tanh_30in_all	34	37	40	33	30
8.	20_10 hid tanh_30in_all	34	37	13	38	37
9.	5_3_2 hid tanh_30in_all	34	26	13	49	30
10.	20_10_2 hid tanh_30in_all	33	37	20	36	33

**Tabela 25** Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem przedziałowym (4 klasy; granica klasy – 35 tygodni) – najlepsze uzyskane wyniki.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas			
			4	3	2	1
1.	10 hid sigma_30in_all	67	100	20	43	77
2.	5 hid sigma_30in_all	63	100	20	30	76
3.	3_2 hid sigma_30in_all	62	100	0	13	82
4.	60_20hid sigma_30in_all	61	100	0	22	77
5.	10_7 hid sigma_30in_all	60	100	0	9	80
6.	5_3 hid fsr_30in_all	61	0	40	26	75
7.	3_2 hid fsr_30in_all	66	0	20	26	83
8.	20_18_10hid tanh_30in_all	65	0	20	26	82
9.	5 hid fsr_30in_all	65	0	20	4	89
10.	15_5 hid tanh_30in_all	64	0	20	22	82

Kolejny – podzielony na dwie części - etap obejmował doświadczenia na danych binarnych. W pierwszym korzystano ze standardowej metody binaryzacji. Jak przedstawiono w rozdziale 5.1.6.2 testowano dwa progi binaryzacji prostej – 30 oraz 35 tygodni. Wyniki modelowania neuronowego dla poszczególnych z nich przedstawiają poniższe tabele. We wszystkich kolejnych przypadkach, kiedy prezentowane są wyniki modelowania zmiennych zależnych o charakterze binarnym zachowana została jednolita notacja wyników. Zapis GLOBAL koduje odsetek wszystkich odpowiedzi prawidłowych bez względu na przynależność do danej klasy, 1 – odsetek prawidłowo sklasyfikowanych rekordów o wartości powyżej przyjętego progu, 0 – odsetek prawidłowo sklasyfikowanych rekordów o wartości poniżej przyjętego progu.

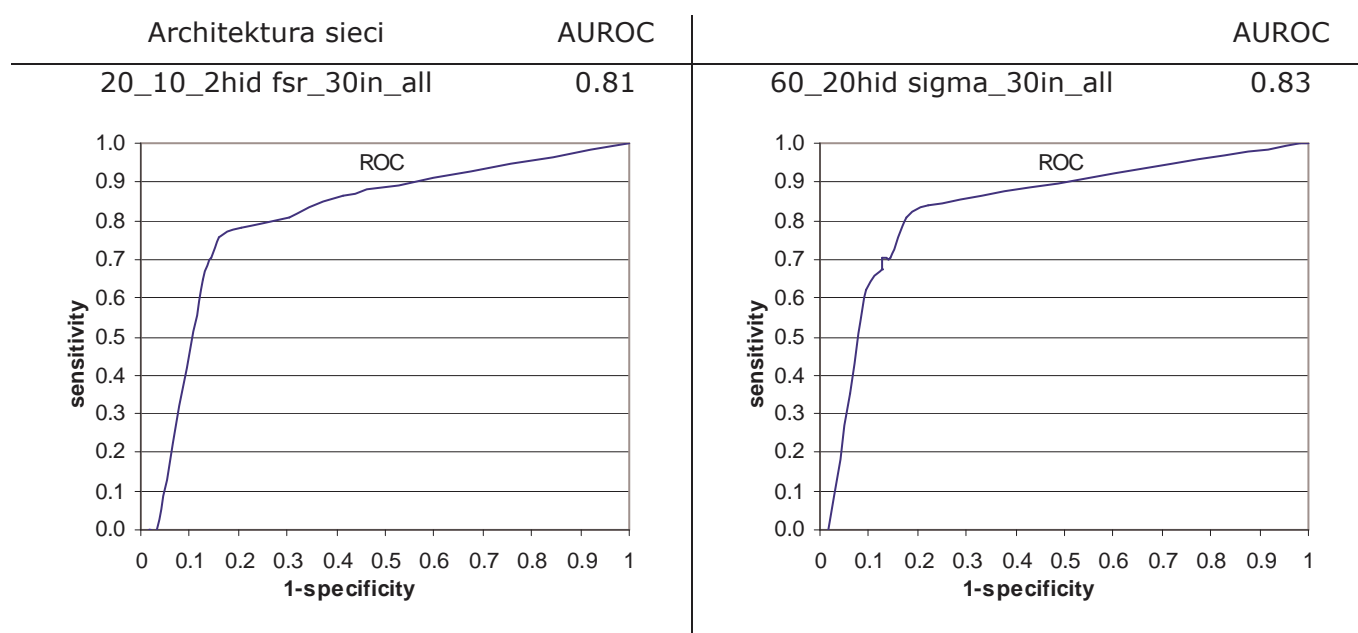


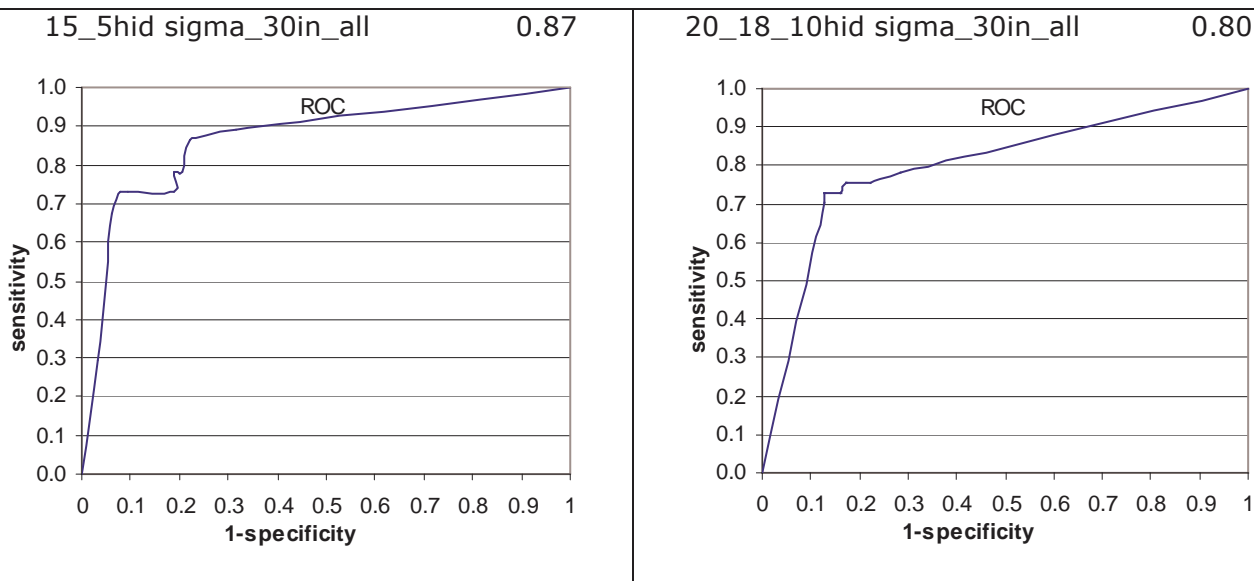
**Tabela 26** Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem binarnym (próg 35 tygodni) – najlepsze uzyskane wyniki.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	20_10_2hid fsr_30in_all	81	76	84	0.81
2.	60_20hid sigma_30in_all	81	70	87	0.83
3.	20_16_8_4hid sigma_30in_all	80	70	86	0.77
4.	20_18_10hid sigma_30in_all	80	73	84	0.80
5.	12_9_7_5_2hid sigma_30in_all	79	68	86	0.76
6.	15_5hid sigma_30in_all	79	76	81	0.87
7.	60_20hid fsr_30in_all	79	73	83	0.78
8.	60_20hid tanh_30in_all	79	65	87	0.77
9.	10_7hid sigma_30in_all	78	73	81	0.79
10.	20_10hid sigma_30in_all	78	70	83	0.80

Poniżej przedstawiono wykresy ROC dla najlepszych modeli uzyskanych na opisywanym etapie (Schemat 30).

**Schemat 30** Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na danych o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść.





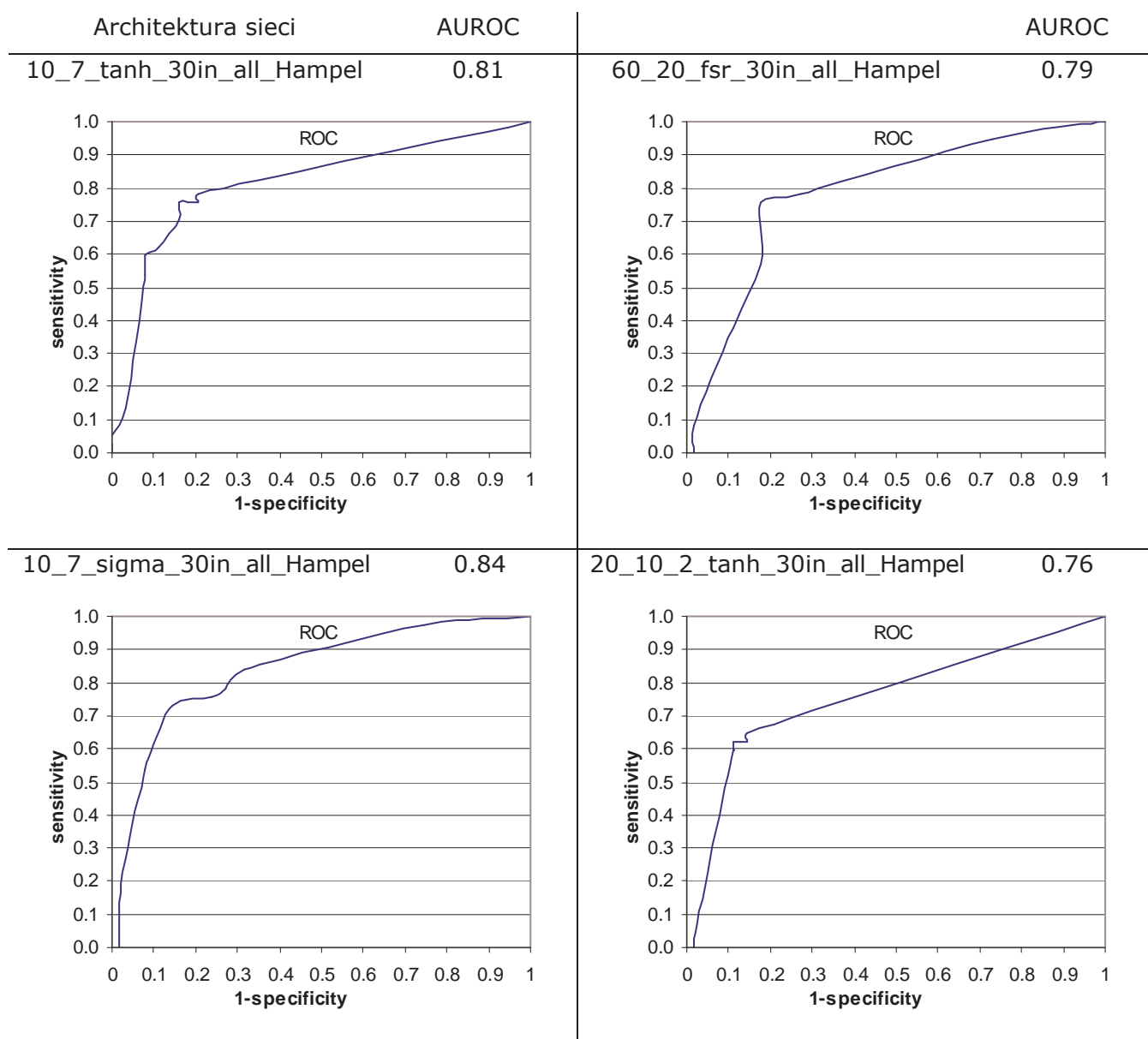
Zmiana funkcji błędu (testowano funkcję Hampela) okazała się nie mieć wpływu na uzyskiwane wyniki. Tabela 27 przedstawia najlepsze uzyskane z opisaną modyfikacją rezultaty.

**Tabela 27** Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem binarnym (próg 35 tygodni) oraz funkcją Hampela jako funkcją kosztu – najlepsze uzyskane wyniki.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	10_7_tanh_30in_all_Hampel	80	73	84	0.81
2.	60_20_fsr_30in_all_Hampel	80	76	83	0.79
3.	10_7_sigma_30in_all_Hampel	78	73	81	0.84
4.	20_10_2_tanh_30in_all_Hampel	78	62	87	0.76
5.	20_10_fsr_30in_all_Hampel	78	70	83	0.77
6.	5_3_fsr_30in_all_Hampel	78	70	83	0.77
7.	60_20_5_tanh_30in_all_Hampel	78	68	84	0.79
8.	18_8_tanh_30in_all_Hampel	77	73	79	0.77
9.	12_9_7_5_2_sigma_30in_Hampel	76	62	84	0.75
10.	15_5_sigma_30in_all_Hampel	76	70	79	0.78

Poniżej przedstawiono wykresy ROC dla najlepszych modeli uzyskanych na opisywanym etapie (Schemat 31).

**Schemat 31** Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na danych o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść z funkcja Hampela jako funkcją kosztu.

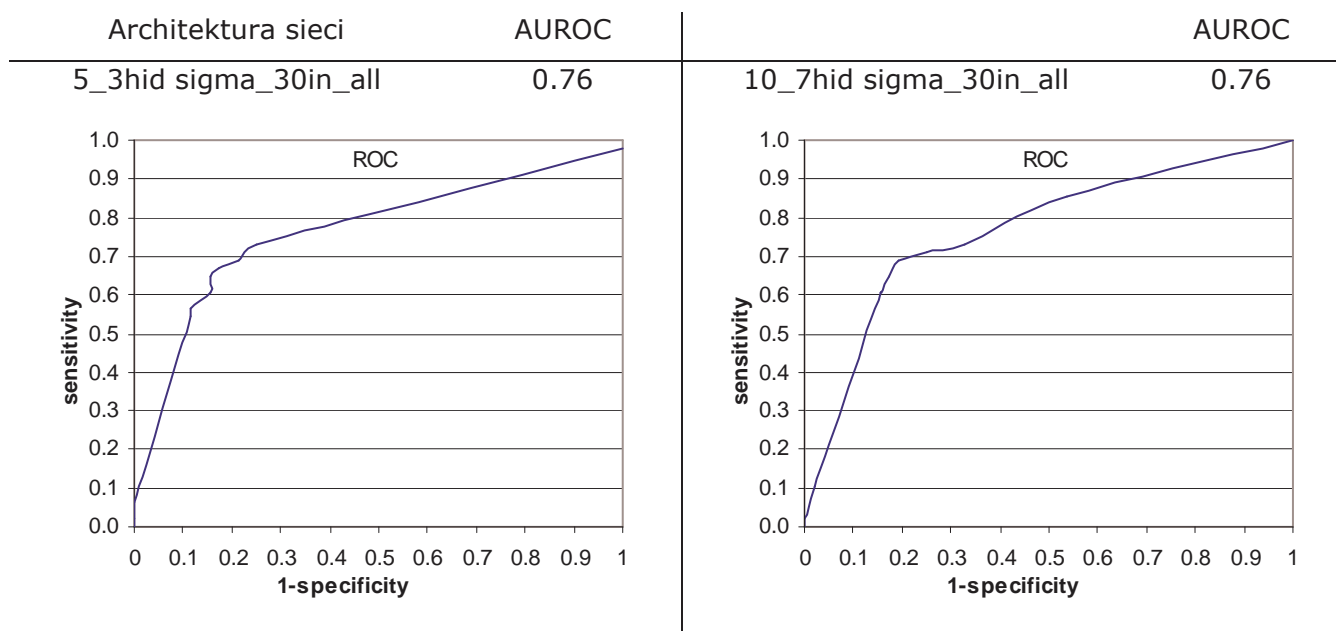


Cechą charakterystyczną zbiorów uzyskiwanych w procesie binaryzacji z zastosowaniem progu o wartości 35-ciu tygodni jest nierówny rozkład klas. Tego typu charakterystyka może negatywnie wpływać na uzyskiwane wyniki. Aby określić skutki przedstawionej asymetrii testowano kolejną modyfikację danych, polegającą na takim doborze wartości progu, który pozwoliłby na wyrównanie liczebności grup. Testowano wartość 30 tygodni.

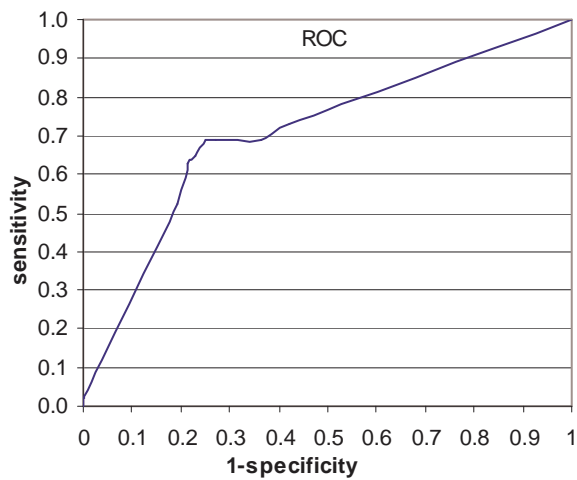
**Tabela 28** Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem binarnym (próg 30 tygodni) – najlepsze uzyskane wyniki.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	5_3hid sigma_30in_all	75	65	85	0.76
2.	10_7hid sigma_30in_all	74	67	81	0.76
3.	18_8hid tanh_30in_all	72	69	75	0.71
4.	20_10_2hid sigma_30in_all	72	63	81	0.74
5.	20_10_2hid fsr_30in_all	71	69	73	0.71
6.	15_5hid sigma_30in_all	70	69	71	0.74
7.	60_20hid fsr_30in_all	70	77	63	0.69
8.	60_20hid tanh_30in_all	70	60	79	0.74
9.	7_5_3hid fsr_30in_all	70	67	73	0.73
10.	7_5_3hid sigma_30in_all	70	65	75	0.75

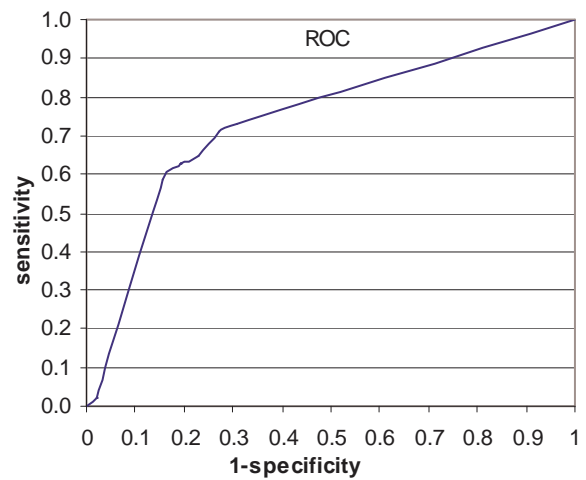
**Schemat 32** Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na danych o binarnym charakterze wyjścia (30 tygodni) dla 30 wejść.



18\_8hid tanh\_30in\_all 0.71



20\_10\_2hid sigma\_30in\_all 0.74



Niezadowalające rezultaty modeli uzyskiwanych w wyniku prowadzonych doświadczeń były powodem kolejnej modyfikacji danych uczących. Bazując na doświadczeniach z etapu modelowania zmiennych o charakterze ciągłym, przeprowadzono analogicznie operacje zaszumiania danych z różną charakterystyką losowo dodawanego szumu.

**Schemat 33** Etap IV – Modelowanie z wykorzystaniem danych przeformatowanych – wyjście binarne, 30 wejść, zaszumienie danych.

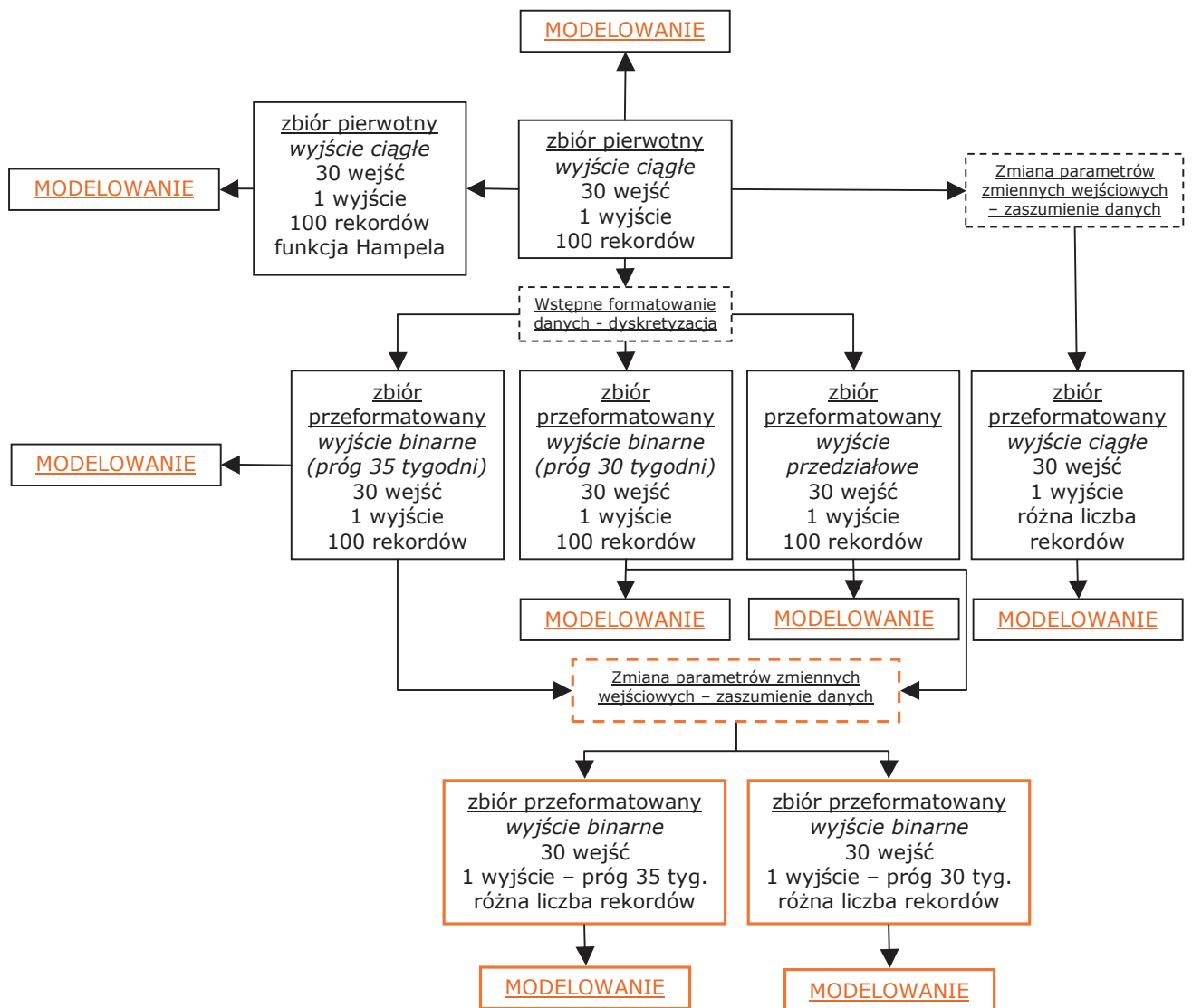


Tabela 29 przedstawia charakterystykę losowego szumu dodawanego na tym etapie do danych pierwotnych oraz ilość rekordów w poszczególnych zbiorach uczących.

**Tabela 29** Parametry zaszumiania danych; ilość rekordów zbiorów wynikowych – wyjście binarne, 30 wejść, próg 35 tygodni.

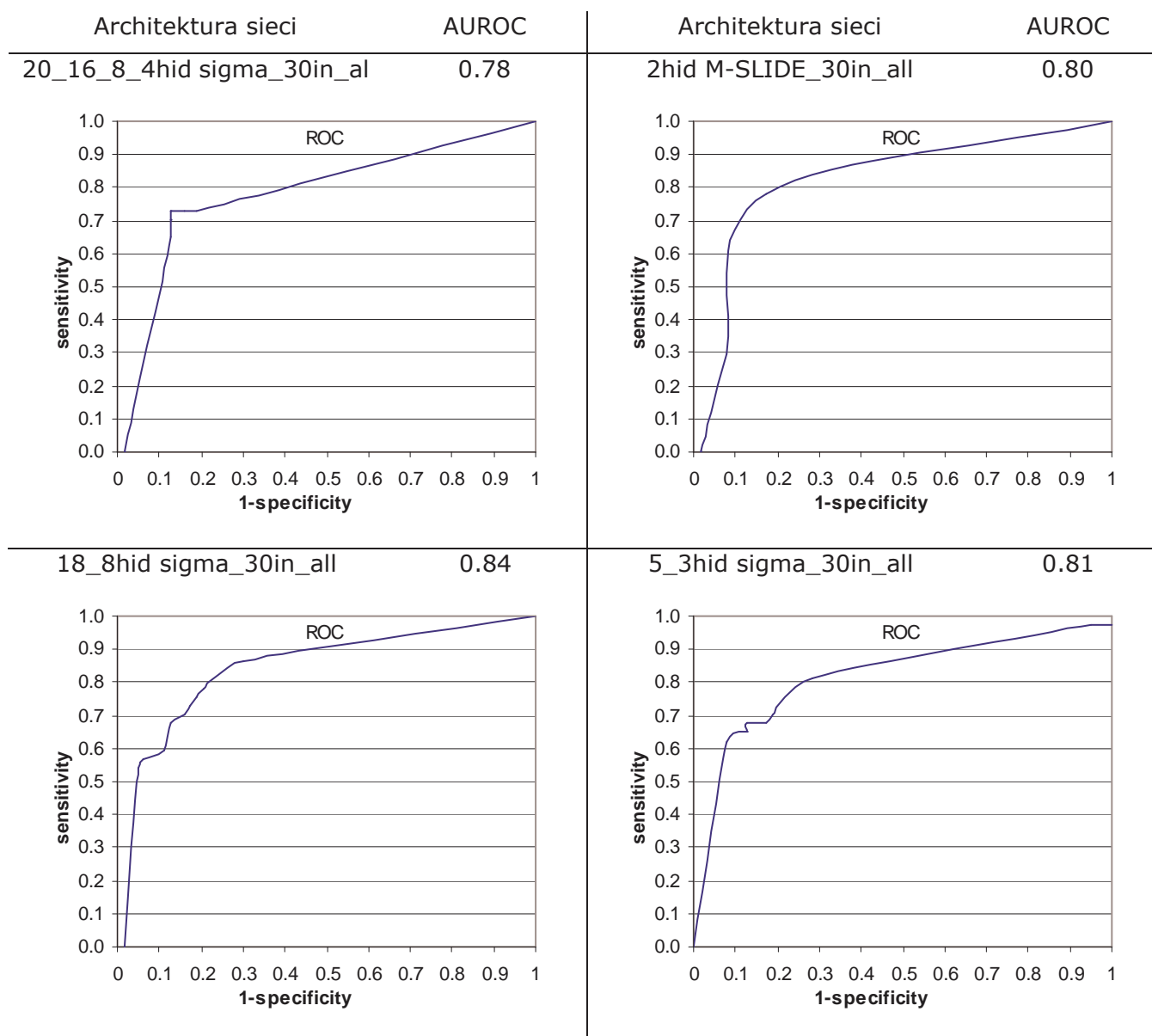
Nr	Opcja zaszumiania	Multiplikacja	Wartość szumu	Ilość rekordów zbioru uczącego
1	szum, randomizacja	2	2%	450
2	szum, randomizacja	2	5%	450
3	szum, randomizacja	2	10%	450
4	szum, randomizacja	5	5%	990
5	szum, randomizacja	10	5%	1890

Poniższe tabele (Tabela 30, Tabela 31, Tabela 32, Tabela 33 oraz Tabela 34) przedstawiają najlepsze wyniki uzyskane na tym etapie doświadczeń, natomiast odpowiednio Schematy 34-38 zawierają wykresy ROC dla kilku najlepszych wybranych modeli.

**Tabela 30** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progów 35 tyg. – najlepsze wyniki dla zbioru o parametrach parametrów szumu 2x2%.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	20_16_8_4hid sigma_30in_all	82	73	87	0.78
2.	2hid M-SLIDE_30in_all	81	70	87	0.80
3.	18_8hid sigma_30in_all	79	70	84	0.84
4.	5_3hid sigma_30in_all	79	65	87	0.81
5.	60_20hid sigma_30in_all	79	68	86	0.81
6.	7_5_3hid sigma_30in_all	79	68	86	0.79
7.	10_7hid sigma_30in_all	78	65	86	0.82
8.	10hid fsr_30in_all	78	70	83	0.79
9.	15_5hid sigma_30in_all	78	68	84	0.81
10.	20_10_2hid sigma_30in_all	78	68	84	0.78

**Schemat 34** Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 2x2%) o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść.

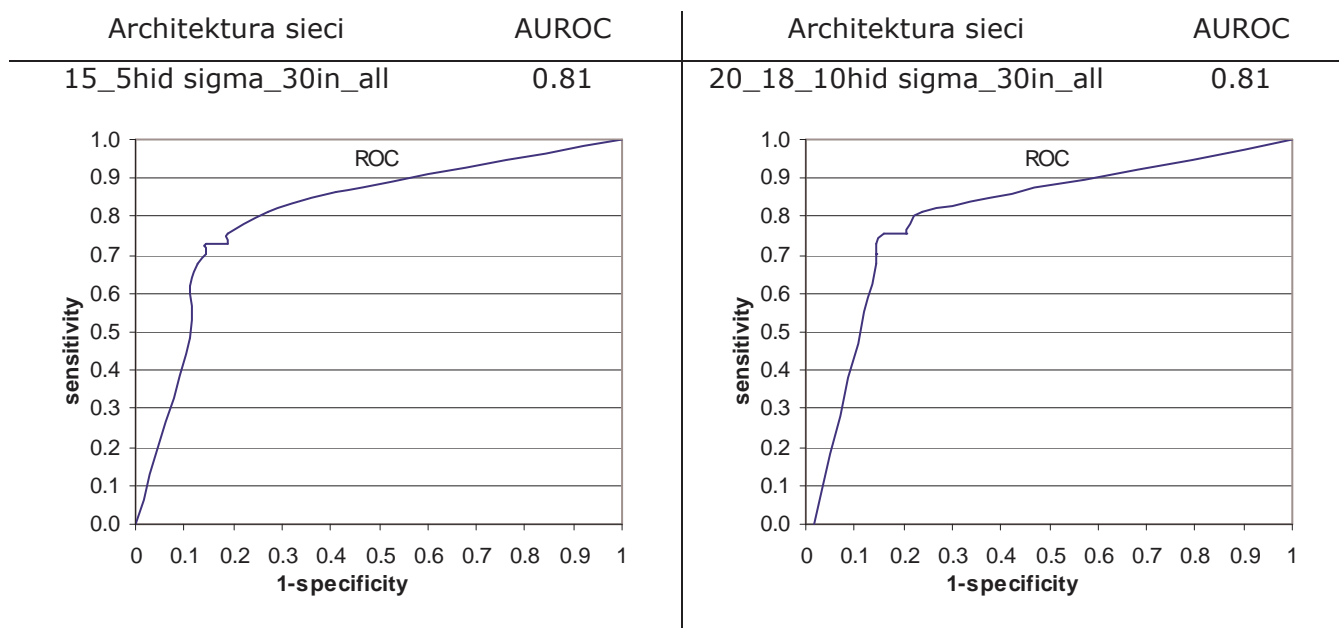


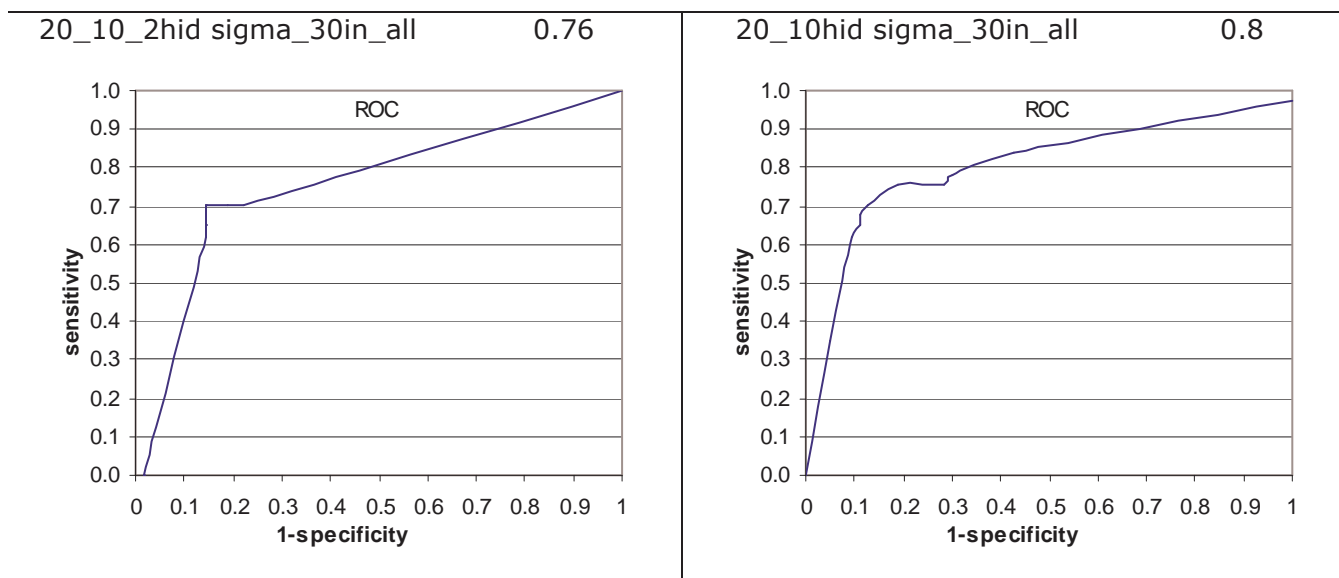


**Tabela 31** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progów 35 tyg. – najlepsze wyniki dla zbioru o parametrach parametrów szumu 2x5%.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	15_5hid sigma_30in_all	81	73	86	0.81
2.	20_18_10hid sigma_30in_all	81	73	86	0.81
3.	20_10_2hid sigma_30in_all	80	70	86	0.76
4.	20_10hid sigma_30in_all	80	70	86	0.8
5.	7_5_3hid sigma_30in_all	80	70	86	0.82
6.	18_8hid sigma_30in_all	79	70	84	0.78
7.	5_3hid sigma_30in_all	79	68	86	0.81
8.	7_5_3hid fsr_30in_all	79	70	84	0.79
9.	10_7hid sigma_30in_all	78	65	86	0.77
10.	10_7hid tanh_30in_all	78	76	79	0.81

**Schemat 35** Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 2x5%) o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść.

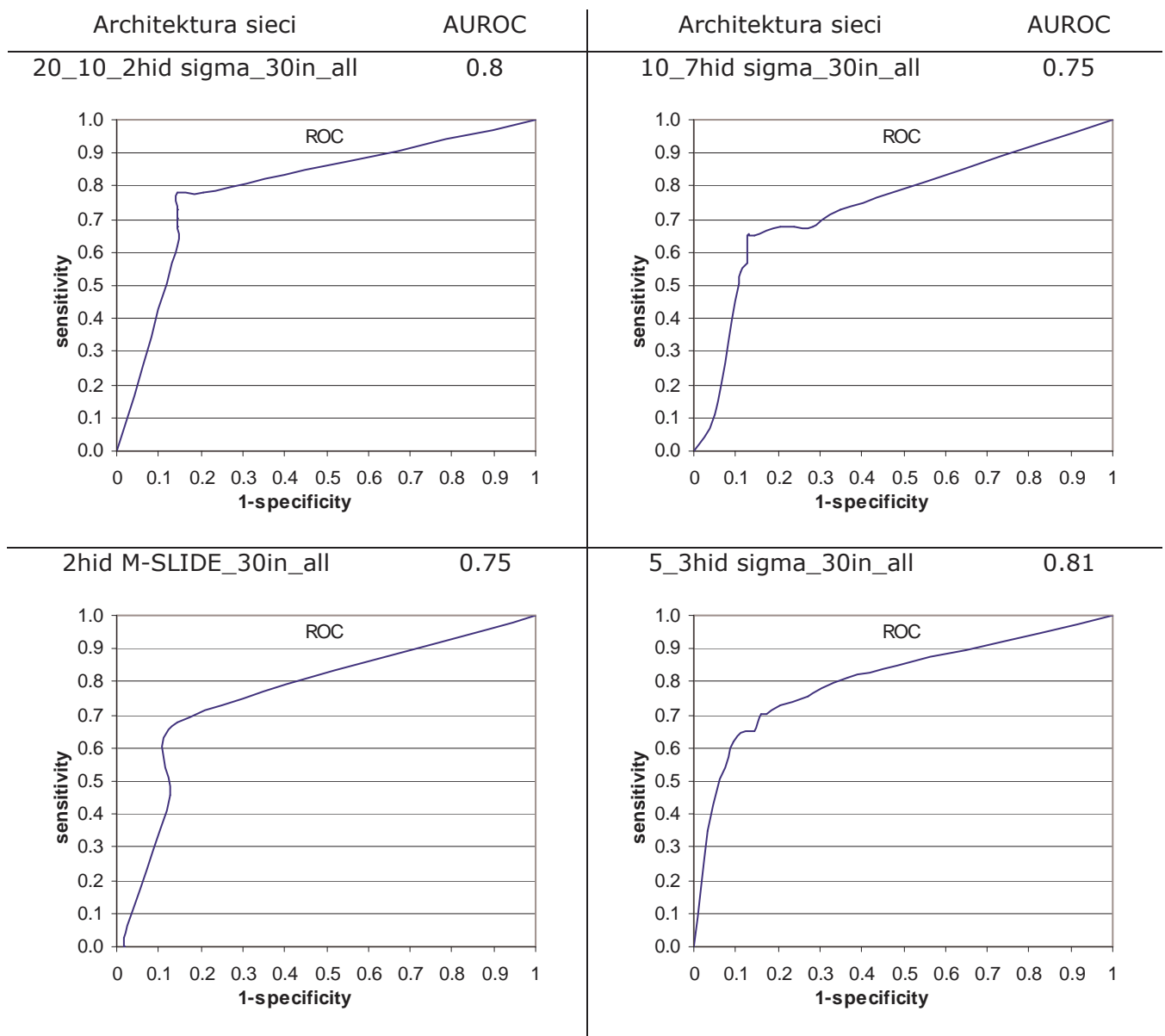




**Tabela 32** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progu 35 tyg. – najlepsze wyniki dla zbioru o parametrach szumu 2x10%.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	20_10_2hid sigma_30in_all	80	70	86	0.8
2.	10_7hid sigma_30in_all	79	68	86	0.75
3.	2hid M-SLIDE_30in_all	79	68	86	0.75
4.	5_3hid sigma_30in_all	79	70	84	0.81
5.	60_20hid sigma_30in_all	79	68	86	0.81
6.	15_5hid fsr_30in_all	78	73	81	0.81
7.	15_5hid sigma_30in_all	78	70	83	0.78
8.	20_18_10hid sigma_30in_all	78	73	81	0.81
9.	5_3hid tanh_30in_all	78	62	87	0.82
10.	10_7hid fsr_30in_all	77	76	78	0.78

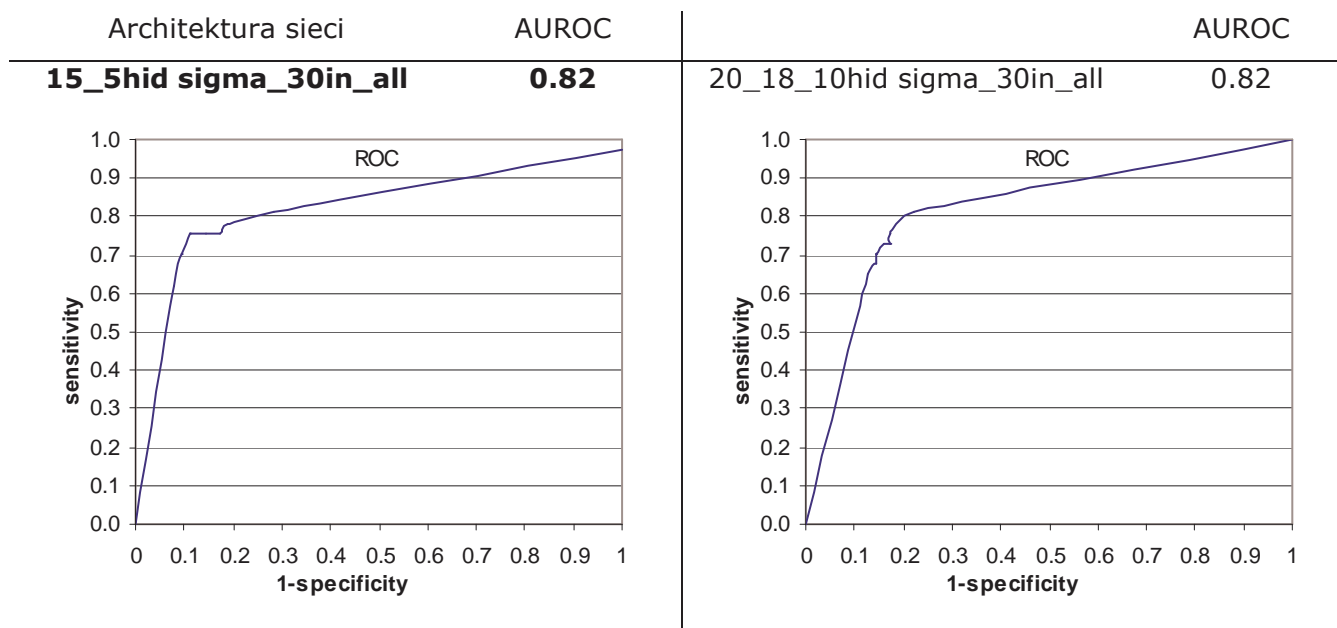
**Schemat 36** Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 2x10%) o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść.

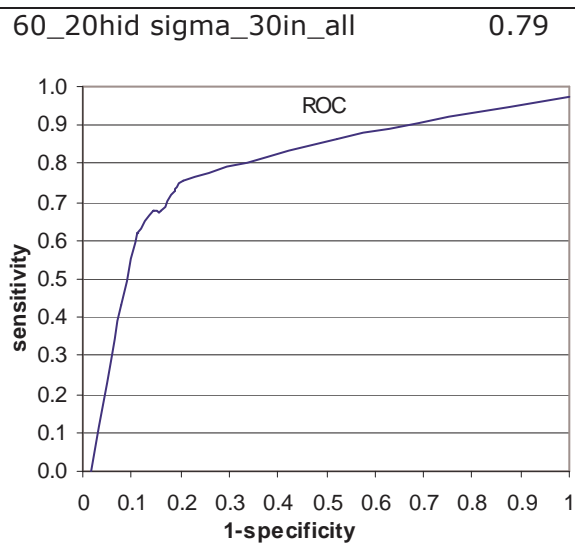
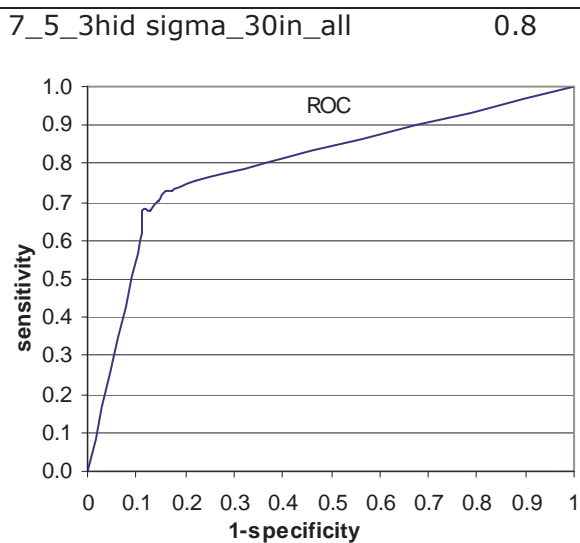


**Tabela 33** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progów 35 tyg. – najlepsze wyniki dla zbioru o parametrach szumu 5x5%.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	<b>15_5hid sigma_30in_all</b>	<b>84</b>	<b>76</b>	<b>89</b>	<b>0.82</b>
2.	20_18_10hid sigma_30in_all	80	73	84	0.82
3.	7_5_3hid sigma_30in_all	80	70	86	0.80
4.	60_20hid sigma_30in_all	79	68	86	0.79
5.	10_7hid sigma_30in_all	78	68	84	0.82
6.	10_7hid tanh_30in_all	78	70	83	0.79
7.	5_3hid sigma_30in_all	78	62	87	0.78
8.	18_8hid sigma_30in_all	77	68	83	0.79
9.	20_10_2hid sigma_30in_all	77	68	83	0.79
10.	20_10hid sigma_30in_all	77	70	81	0.80

**Schemat 37** Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 5x5%) o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść.

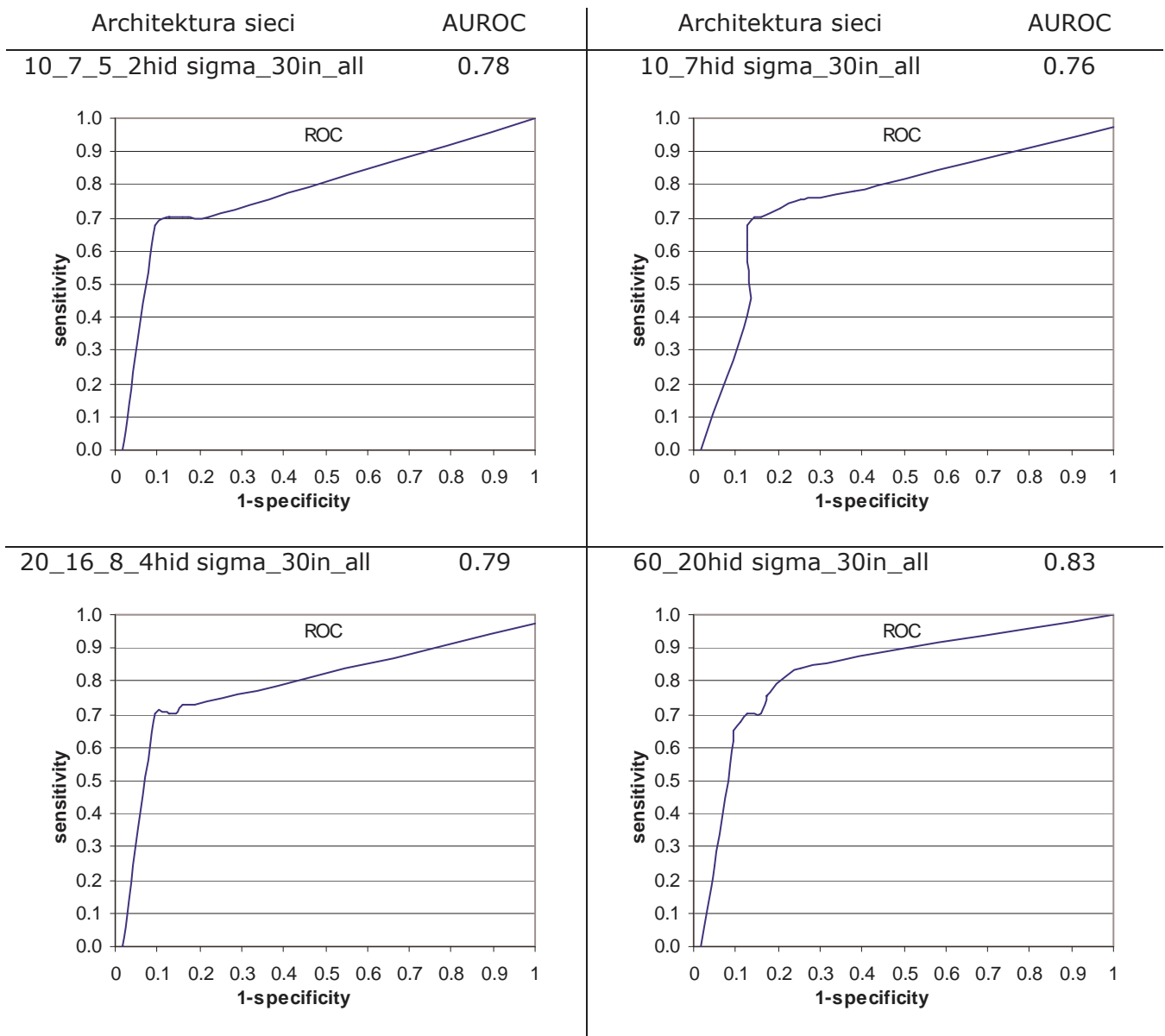




**Tabela 34** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progu 35 tyg. – najlepsze wyniki dla zbioru o parametrach parametrów szumu 10x5%.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	10_7_5_2hid sigma_30in_all	81	70	87	0.78
2.	10_7hid sigma_30in_all	80	70	86	0.76
3.	20_16_8_4hid sigma_30in_all	80	70	86	0.79
4.	60_20hid sigma_30in_all	80	70	86	0.83
5.	15_5hid sigma_30in_all	79	70	84	0.84
6.	20_10_2hid sigma_30in_all	79	70	84	0.80
7.	20_10_2hid tanh_30in_all	79	68	86	0.78
8.	20_18_10hid sigma_30in_all	79	73	83	0.82
9.	2hid M-SLIDE_30in_all	79	65	87	0.79
10.	40_fsr_40in_all	79	68	86	0.78

**Schemat 38** Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 10x5%) o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść.

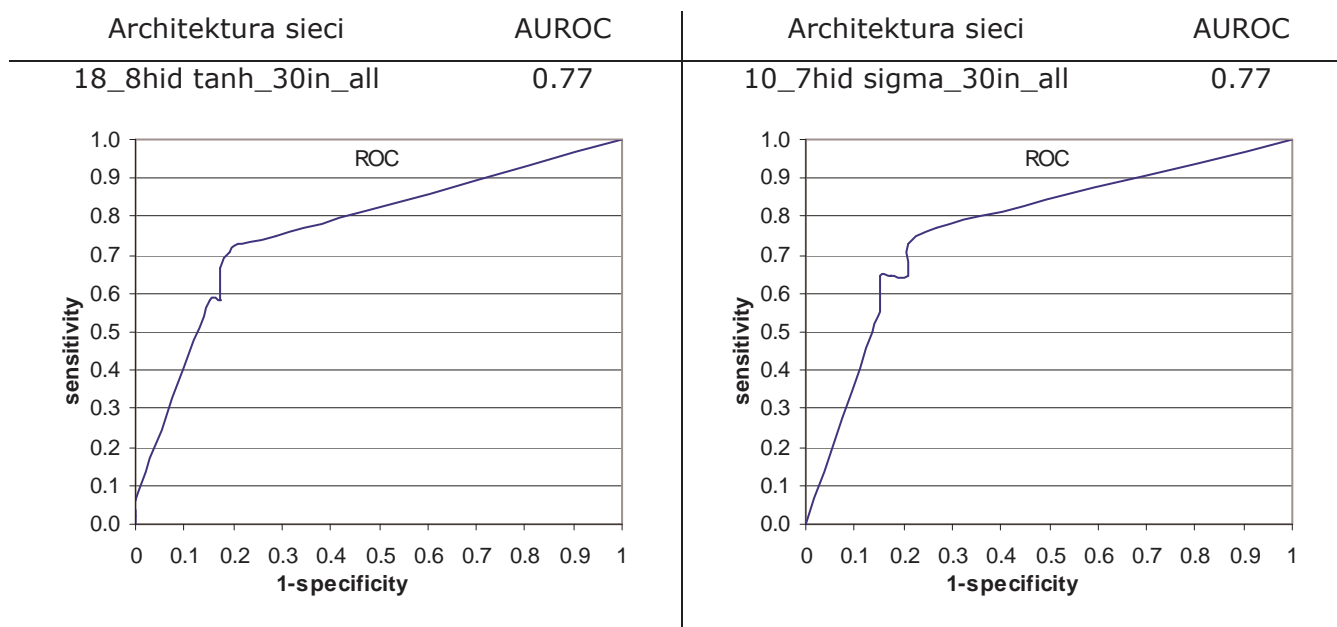


Ze względu na gorsze wyniki uzyskiwane w trakcie modelowania z wykorzystaniem zbiorów binaryzowanych w oparciu o 30-to tygodniowy próg przeżycia, przetestowano jedynie jeden schemat zaszumiania danych. Wybrano ten, z wykorzystaniem którego uzyskano najlepsze wyniki w trakcie testowania proggu 35-cio tygodniowego (5x5%). Tabela 35 oraz Schemat 39 zawierają najlepsze uzyskane wyniki oraz ich zobrazowanie przy pomocy krzywej ROC.

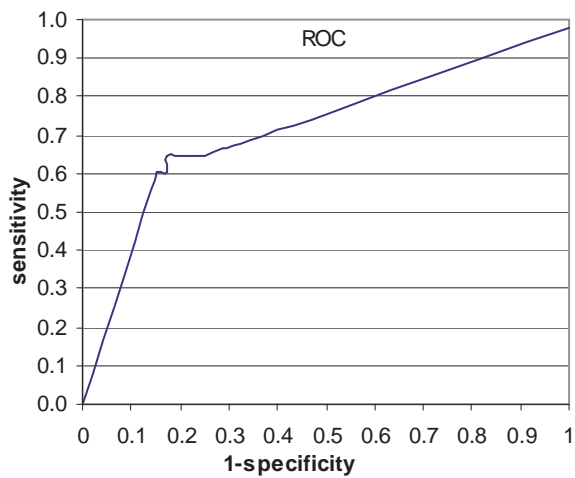
**Tabela 35** Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progu 30 tyg. – najlepsze wyniki dla zbioru o parametrach parametrów szumu 5x5%.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	18_8hid tanh_30in_all	75	67	83	0.77
2.	10_7hid sigma_30in_all	74	65	83	0.77
3.	15_5hid tanh_30in_all	74	65	83	0.72
4.	18_8hid sigma_30in_all	74	69	79	0.77
5.	7_5_3hid sigma_30in_all	73	67	79	0.74
6.	20_10hid sigma_30in_all	72	65	79	0.73
7.	20_10_2hid tanh_30in_all	71	73	69	0.73
8.	20_18_10hid sigma_30in_all	71	60	81	0.72
9.	60_20hid sigma_30in_all	71	69	73	0.75
10.	15_5hid fsr_30in_all	70	65	75	0.73

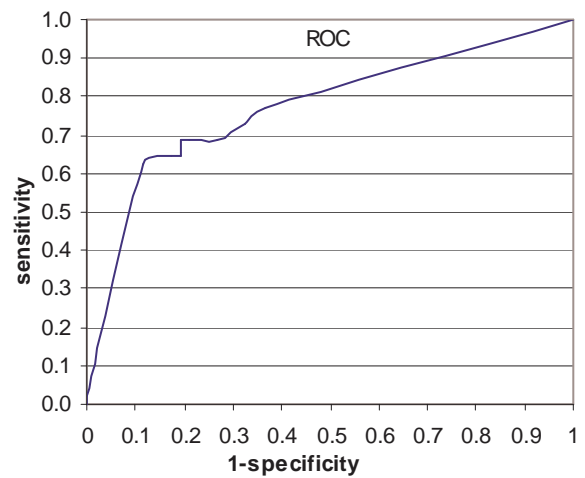
**Schemat 39** Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 5x5%) o binarnym charakterze wyjścia (30 tygodni) dla 30 wejść.



15\_5hid tanh\_30in\_all 0.72



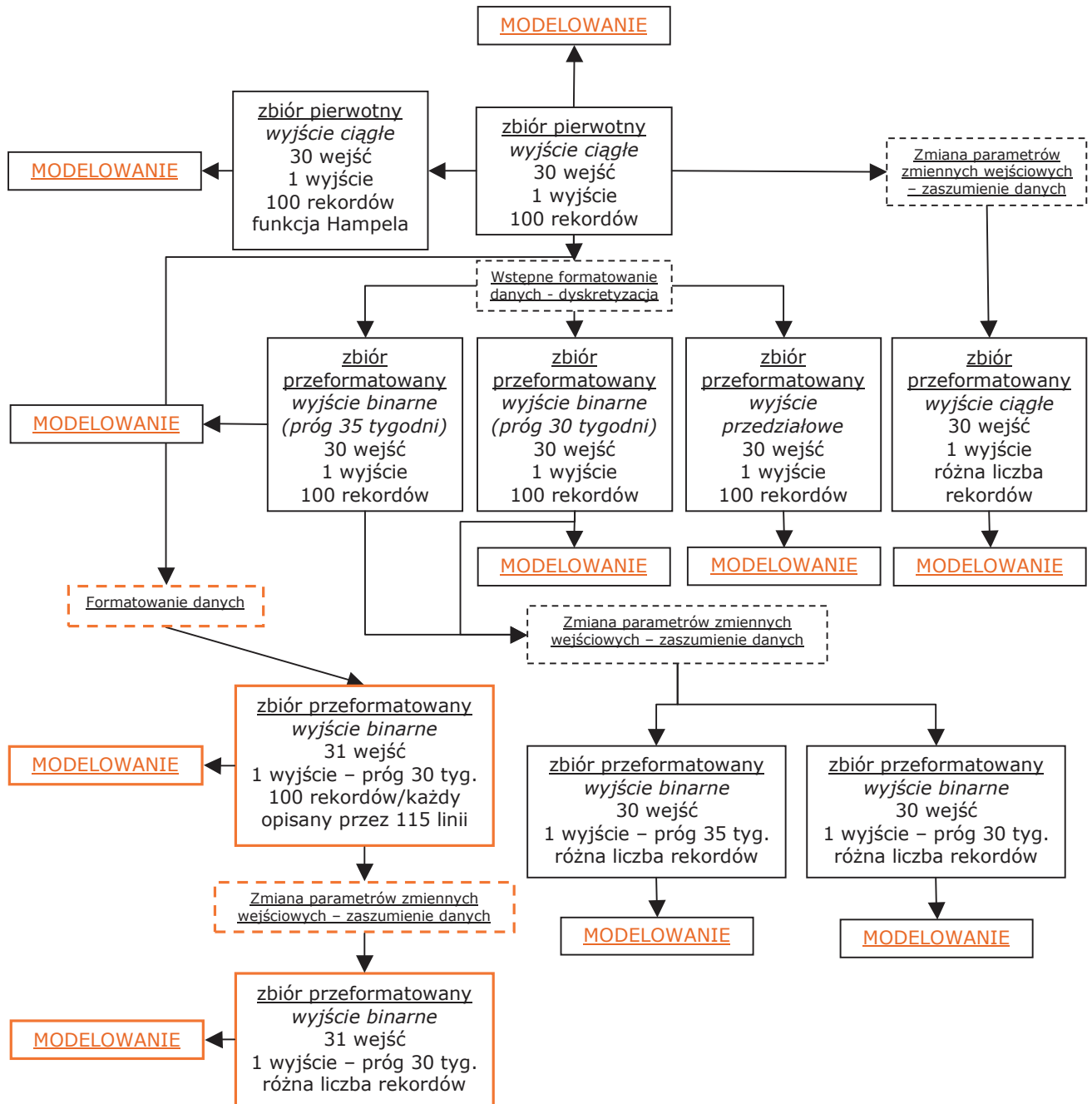
18\_8hid sigma\_30in\_all 0.77



Kolejne doświadczenie obejmowało zarówno zmianę charakteru wyjścia (binarne), jak również modyfikację sposobu podawania informacji wejściowej. Powyższa zmiana wynikała z odmiennego od standardowego sposobu binaryzacji, a wynikiem jego zastosowania była zmiana rozmiaru wektora wejściowego (zmiana z 30 do 31 wejść). Szczegółowe omówienie metodyki uzyskiwania par zbiorów uczący – testowy przedstawiony został w rozdziale 5.1.6.2.



**Schemat 40** Etap V – Modelowanie z wykorzystaniem danych przeformatowanych – wyjście binarne, 31 wejść, zaszumienie danych.

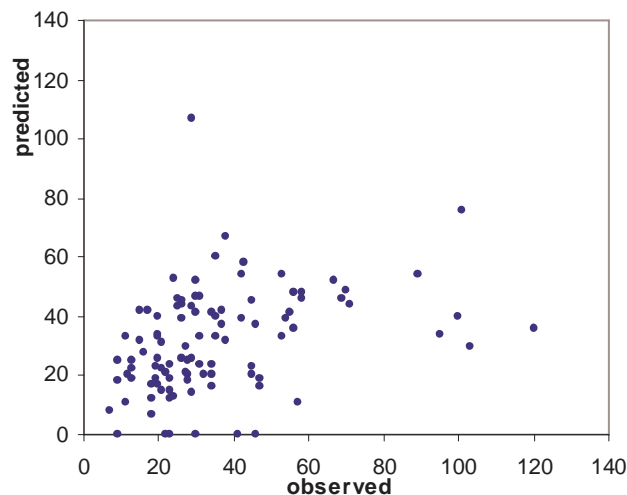


Na obecnym etapie zastosowano dwa schematy testowania skuteczności uzyskanych modeli neuronowych. Pierwszy z nich obejmował klasyczny podział na pary testowy-uczący oraz ich dalszą obróbkę w celu uzyskania wektora o 31 wejściach. Zbiór testowy zawierał 10% rekordów zbioru pierwotnego, a każdy z nich opisany był – zgodnie z przedstawioną powyżej metodyką – przez 115 wierszy. W trakcie opisywanych

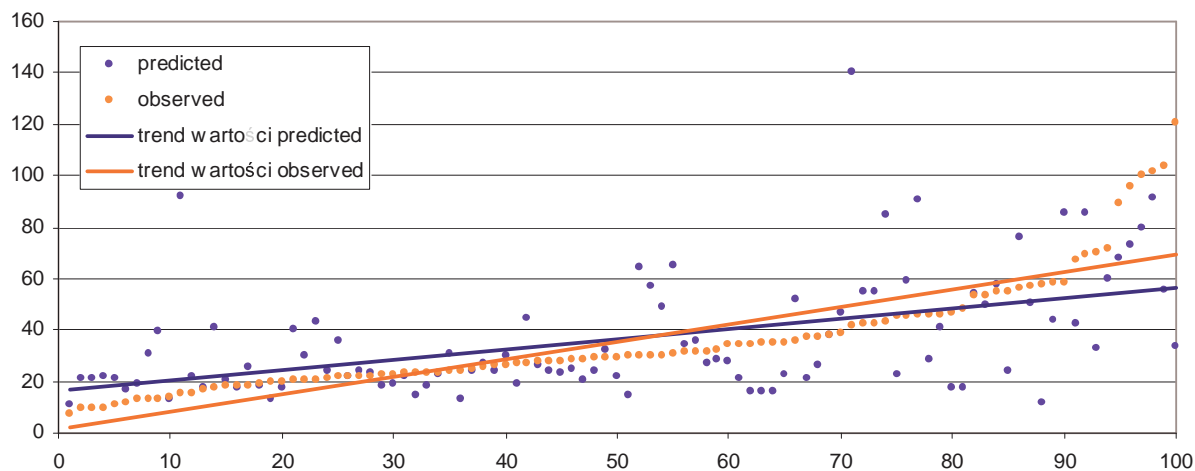
doświadczeń wykonano analizę wpływu zmian wartości granicznego progu pobudzenia na uzyskiwane wyniki generalizacji. Modyfikacja polegała na obliczeniu średniej wartości pobudzenia (średnia obliczona z wartości minimalnej i maksymalnej) dla każdego rekordu. W kolejnym kroku rekordy zbiorów testowych przedstawione zostały w postaci dwóch wierszy – zakodowano jedynie wartości graniczne.

Tabela 36 przedstawia najlepsze uzyskane wyniki dla domyślnej wartości progu pobudzenia (0,5), natomiast Tabela 37 przedstawia wyniki analogicznych architektur dla modyfikowanej wartości progu pobudzenia.

**Schemat 41** Wykres zależności wartości rzeczywistej do przewidzianej przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z domyślną wartością progu.



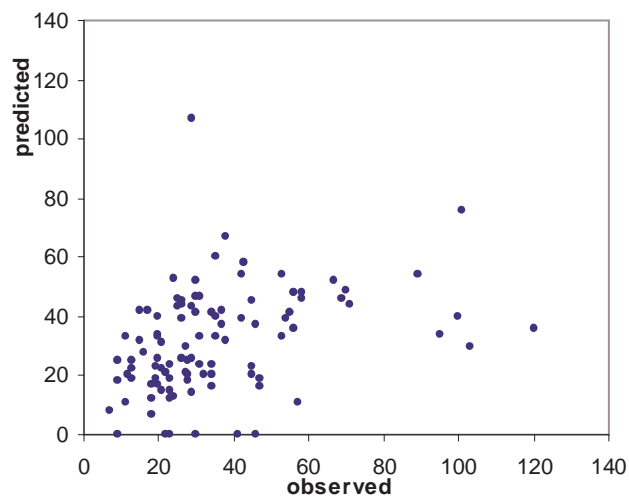
**Schemat 42** Porównanie wartości rzeczywistych i przewidzianych przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z domyślną wartością proggu; analiza trendów



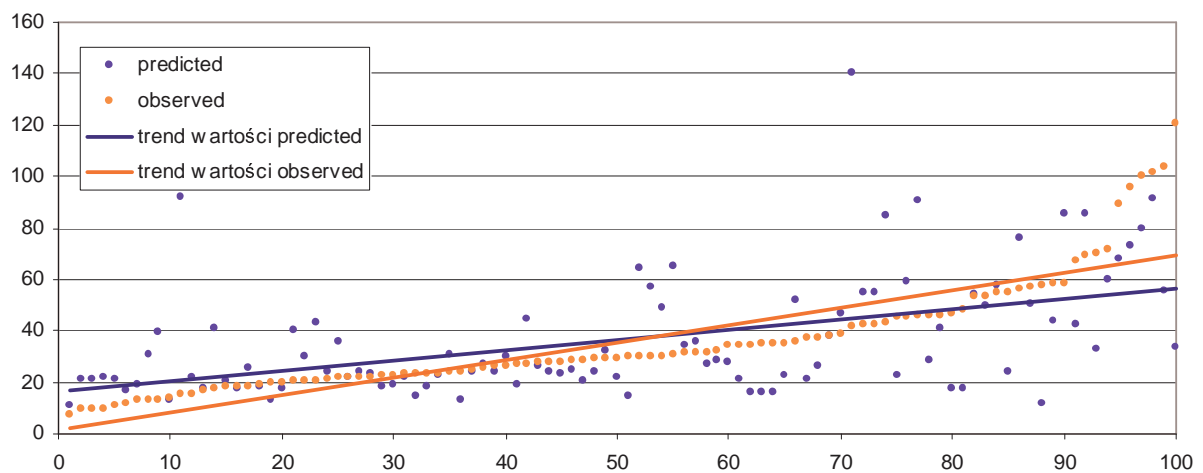
**Tabela 36** Architektury testowane w trakcie doświadczeń na danych z wyjściem binarnym, opisanych przez 31 wejść dla domyślnej wartości proggu - najlepsze wyniki.

Nr	Architektura sieci	Wartość RMSE
1.	2hid fsr_31in_all	23.10
2.	2hid sigma_31in_all	23.74
3.	10_7hid tanh_31in_all	23.91
4.	100hid sigma_31in_all	23.94
5.	100hid fsr_31in_all	24.30
6.	30hid sigma_31in_all	24.47
7.	30hid fsr_31in_all	24.53
8.	5_3hid fsr_31in_all	24.70
9.	20_10hid fsr_31in_all	24.89
10.	10_7hid sigma_31in_all	24.92

**Schemat 43** Wykres zależności wartości rzeczywistej do przewidzianej przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z domyślną wartością progów.



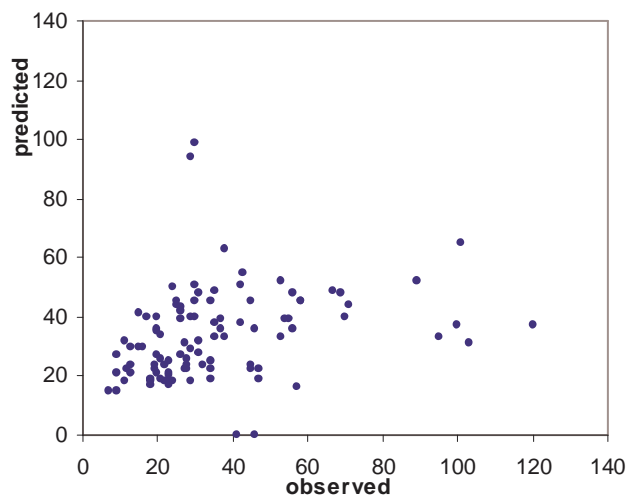
**Schemat 44** Porównanie wartości rzeczywistych i przewidzianych przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z domyślną wartością progów; analiza trendów



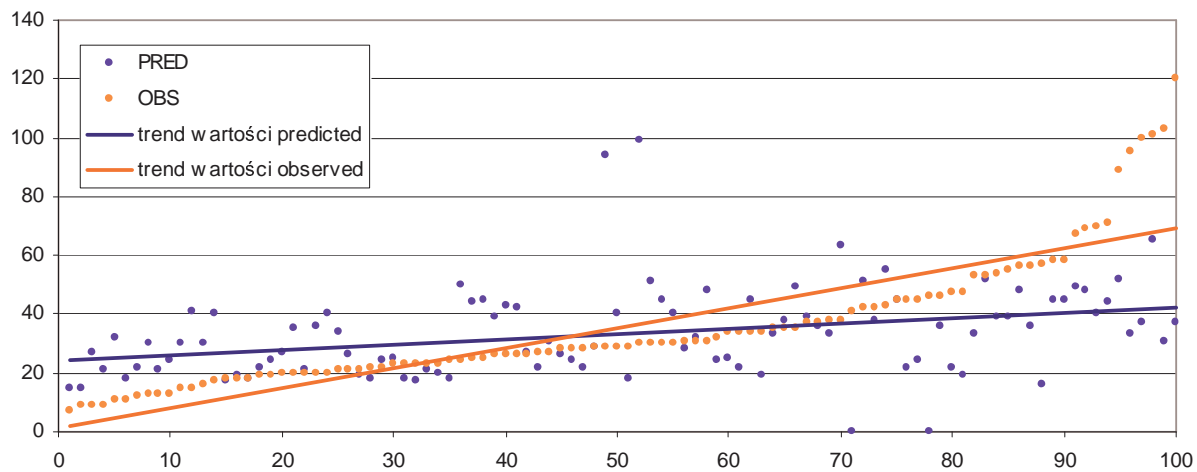
**Tabela 37** Architektury testowane w trakcie doświadczeń na danych z wyjściem binarnym, opisanych przez 31 wejść dla modyfikowanej wartości progu - najlepsze wyniki.

Nr	Architektura sieci	Wartość RMSE
1.	2hid fsr_31in_all	23.11
2.	100hid sigma_31in_all	23.71
3.	30hid fsr_31in_all	24.18
4.	10_7hid tanh_31in_all	24.19
5.	100hid fsr_31in_all	24.30
6.	10hid fsr_31in_all	24.36
7.	5hid fsr_31in_all	24.57
8.	30hid sigma_31in_all	24.65
9.	20_10hid fsr_31in_all	25.13
10.	5_3hid fsr_31in_all	25.71

**Schemat 45** Wykres zależności wartości rzeczywistej do przewidzianej przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z modyfikowaną wartością progu.



**Schemat 46** Porównanie wartości rzeczywistych i przewidzianych przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z domyślną wartością proggu; analiza trendów



## 6 Zastosowanie sztucznych sieci neuronowych w epidemiologii

Epidemiologia jest jedną z dziedzin nauk medycznych, w której najtrudniej zastosować modele matematyczne w celu określenia odległych skutków zmiany pewnych parametrów (fizjologicznych np. poziom cholesterolu, demograficznych np. śmiertelność). Głównymi powodami takiego stanu rzeczy są wielowymiarowość i nieliniowość modelowanych zjawisk oraz konieczność dokonywania ekstrapolacji poza posiadane dane. Jednocześnie tego typu modele są powszechnie stosowane, niekiedy w codziennej praktyce lekarskiej [64], jednak najbardziej znanym przykładem takiego modelu epidemiologicznego, najczęściej cytowanym i powszechnie wykorzystywanym - jest model regresyjny ryzyka choroby wieńcowej serca, oparty na danych z badania Framingham Heart Study [64]. W oparciu o informacje z tego - kontynuowanego od 1948 roku - badania, budowane są również inne modele i określane czynniki ryzyka innych zespołów chorobowych [27,79]. Przyczyną dla której podobne systemy są konstruowane, jest ich duża przydatność praktyczna oraz zdolność do wykazania kierunku i wartości odległych zmian skutków ostatecznych na podstawie jedynie czynników ryzyka.

### 6.1 Metodyka

#### 6.1.1 Dane

Zobrazowanie zalet neuronowej analizy dużych zbiorów danych możliwe będzie dzięki wykorzystaniu baz udostępnianych przez National Center for Health Statistics (NCHS), Center for Disease Control and Prevention (CDC), instytucję zajmującą się gromadzeniem, analizą i rozpowszechnieniem informacji dotyczących stanu zdrowia populacji amerykańskiej [61].

Wykorzystany w badaniu zbiór informacji opisujący badanie Behavioral Risk Factor Surveillance System (BRFSS) [61] opisuje losową grupę ok. 150 tysięcy dorosłych respondentów, o których informacje zbierane były drogą telefoniczną. Celem badania było zebranie informacji

dotyczących zachowań zdrowotnych i powiązania ich z występującymi chorobami zakaźnymi i przewlekłymi.

Badanie National Health and Nutrition Examination Survey (NHANES) jest badaniem zaplanowanym i przeprowadzonym w celu określenia stanu zdrowia oraz sposobu odżywiania populacji amerykańskiej. Badanie NHANES I Epidemiologic Followup Study (NHEFS) jest kontynuacją badania NHANES (National Health and Nutrition Examination Survey). W bazie danych znajduje się podstawowy opis stanu zdrowia ok. 15 tysięcy pacjentów. Informacje zbierane były drogą telefoniczną. Opis poszczególnych zbiorów danych przedstawiony został w Tabeli 38.

**Tabela 38** Epidemiologiczne zbiory danych wykorzystane w pracy

L.p.	Nazwa	Wykorzystany podzbiór informacji	Liczba respondentów	Liczba zmiennych
1.	Behavioral Risk Factor Surveillance System (BRFSS_1999)	-	159 989	284
2.	National Health and Nutrition Examination Survey (NHANES)	NHANES I Epidemiologic Follow up Study (NHEFS)	11 195	1844

Ze względu na rozmiar zbiorów pierwotnych (liczba zmiennych sterujących) ich dokładna charakterystyka została przedstawiona w postaci załącznika do niniejszego opracowania - Załącznik 1.

#### 6.1.2. Wartości wyjściowe (modelowane)

Przedstawione powyżej zbiory wykorzystywane w trakcie prac różnią się konstrukcją i sposobem kodowania zmiennych. W związku z tym każdy problem wymaga osobnego przygotowania danych uczących oraz informacji podlegającej modelowaniu.

Schemat 44 przedstawia obszar zagadnień wykorzystanych do zbudowania drogi wnioskowania potwierdzającej możliwość zastosowania

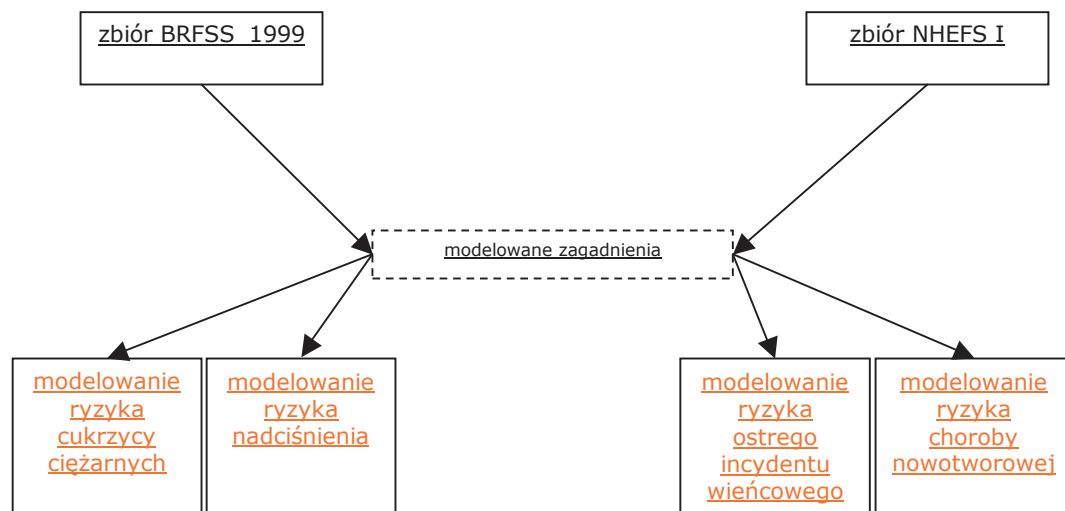


sztucznych sieci neuronowych jako narzędzi modelowania w epidemiologii.

Podstawowe problemy to:

- modelowanie ryzyka wystąpienia nadciśnienia
- modelowanie ryzyka ostrego incydentu wieńcowego (ataku serca)
- modelowanie ryzyka wystąpienia choroby nowotworowej
- modelowanie ryzyka wystąpienia cukrzycy ciężarnych

**Schemat 47** Schemat zagadnień z zakresu modelowania epidemiologicznego



We wszystkich opisywanych zagadnieniach wartość wyjściowa miała charakter binarny. Zadaniem modelu była prawidłowa klasyfikacja przypadku do klasy przypadków choroby lub wolnych od schorzenia.

### 6.1.3. Preprocessing i zaszumianie danych

Ze względu na szeroki zakres opracowywanych zagadnień i wynikającą z tego specyfikę poszczególnych zbiorów danych ich charakterystyka oraz drogi modyfikacji podawanych sieci informacji są przedstawione osobno dla każdego z działów modelowania neuronowego w epidemiologii.

## 6.2 Rezultaty poszczególnych etapów pracy – modelowanie ryzyka wystąpienia cukrzycy ciężarnych (BRFSS\_1999)

### 6.2.1. Opis problemu

Cukrzyca ciężarnych jest definiowana jako trzeci typ cukrzycy. Zgodnie z najnowszymi wytycznymi jest to każda nietolerancja glukozy rozpoznana w okresie ciąży [108]. Wśród najważniejszych czynników ryzyka wymieniane są [24,108]:

- otyłość
- wiek matki powyżej 40 lat
- wcześniejsze poronienia
- rasa
- styl życia

Schorzenie to dotyczy - w zależności od czynników środowiskowych i genetycznych – od 3 do 12% ciężarnych kobiet. Pomimo, że jedynie w ok. 7% przypadków występują powikłania nietolerancji glukozy, to jednak ich skutki zarówno dla matki jak i przede wszystkim dla dziecka są poważne. Niewyrównana glikemia może skutkować poronieniem w ostatnich dwóch miesiącach ciąży, makrosomią płodu i innymi zaburzeniami wzrostu, hipoglikemią czy żółtaczką noworodka. Niemniej niebezpieczne są powikłania odległe, a w tym: wzrost ryzyka rozwoju cukrzycy II typu matki i dziecka, chorobliwa otyłość [24]. Z tego względu towarzystwa naukowe (American Diabetes Association, Polskie Towarzystwo Diabetologiczne) w swoich wytycznych postulują konieczność skryningu kobiet ciężarnych pod kątem ewentualnej nietolerancji glukozy [62,108]. Jednocześnie jednak ze względu na kosztowność i uciążliwość dla pacjentki poszukuje się metod i schematów, które byłyby efektywne kosztowo i jednocześnie nie powodowałyby negatywnego wpływu na jakość życia kobiet w okresie ciąży. Propozycje obejmują m.in. przeprowadzanie badania jedynie w grupach ryzyka (klasyfikacja w oparciu o prosty model oparty na wartości BMI) [71].

Zastosowanie sztucznych sieci neuronowych jako narzędzi predykcji ryzyka wystąpienia cukrzycy ciężarnych i publikacja odpowiedniego narzędzia za pośrednictwem sieci Internet, pozwala spełnić wymogi postawione na zakończenie poprzedniego akapitu. Założeniem projektu była budowa dostępnego, mało uciążliwego dla pacjenta, skutecznego narzędzia skringowego, bazującego na prostych danych epidemiologicznych.

#### 6.2.2. Przygotowanie danych

Wykorzystany w trakcie doświadczeń zbiór danych (BRFSS\_1999) filtrowano pod kątem przygotowanych kryteriów włączenia do grupy badanej. Rekordy opisujące kobiety zostały przebadane pod następującym kątem:

- ciąża w ostatnich pięciu latach
  - brak zdiagnozowanego innego typu cukrzycy
- Jednocześnie kryteriami wyłączenia ze zbioru bazowego były:
- zdiagnozowany inny typ cukrzycy w okresie przed ciążą
  - brak jakiegokolwiek z informacji w rekordzie danych

Spośród 2551 wyselekcjonowanych pacjentek, u 91 (3,7% przypadków) została stwierdzona cukrzyca ciężarnych. Pierwotnie każdy rekord opisany był 197 zmiennymi. Ze względu na nierówny rozkład przypadków zaklasyfikowanych do oby klas (pozytywna i negatywna), spodziewano się braku elastyczności systemu neuronowego uzyskiwanego w wyniku prac z tak przygotowanym zbiorem danych. W związku z tym zastosowano metodykę pozwalającą na uniknięciu efektu nauczania się przez sieć jedynie jednej, dominującej klasy – balansowanie klas. Dwie odmiany przedstawianej metody określane są jako balansowanie ujemne oraz balansowanie dodatnie.

- metoda balansowania ujemnego – losowo wybrane rekordy z liczniej reprezentowanej klasy są usuwane z plików uczących aż do wyrównania liczebności klas; pliki testowe pozostają bez zmian; efektem balansowania ujemnego jest zmniejszenie liczby rekordów

plików uczących (co niekiedy może skutkować usunięciem rekordów niosących ważną informację) i jednocześnie wyrównanie liczebności klas

- balansowania dodatnie – losowo wybrane rekordy plików uczących, z klasy mniej licznie reprezentowanej są multiplikowane aż do wyrównania liczebności klas; pliki testowe pozostają bez zmian; efektem balansowania dodatniego jest wzrost liczby rekordów plików uczących (jednak bez wprowadzania nowej informacji) i jednocześnie wyrównanie liczebności klas

Balansowania klas dokonano wykorzystując program *Preproc* [86].

Wśród 197 deskryptorów każdego rekordu, na podstawie dostępnej literatury i opiniach ekspertów wyselekcjonowano ręcznie 8 zmiennych [24,51,62,71,108]. Selekcję prowadzono pod kątem:

- wpływu na możliwość rozwinięcia się nietolerancji glukozy w trakcie ciąży (czynniki ryzyka)
- dostępności informacji dla przeciętnego pacjenta – ponieważ jednym z założeń projektu była konstrukcja systemu skutecznego ale jednocześnie powszechnie dostępnego i łatwego w obsłudze nie tylko dla specjalistów

Tabela 39 przedstawia zbiorczo opis danych wykorzystanych w trakcie modelowania ryzyka cukrzycy ciężarnych.

**Tabela 39** Zmienne opisujące rekordy pacjentek włączonych do badania (cukrzyca ciężarnych, 8 wejść).

L.p.	Kod zmiennej	Opis	Możliwe wartości
1.	BPHIGH	Zdiagnozowane nadciśnienie tętnicze	1-tak/0-nie
2.	TOLDHI	Zdiagnozowany podniesiony ponad normę poziom cholesterolu	1-tak/0-nie
3.	SMOKEDAY	Częstotliwość palenia tytoniu	0-nie wiem/1-codziennie/2-od czasu do czasu/3-nie pale
4.	AGE	Wiek	[lata]
5.	ORACE	Rasa	1-kaukaska/2-czarna/3-azjatycka/4-indianie amerykańscy/5-inna
6.	WEIGHT	Waga	[kg]
7.	HEIGHT	Wzrost	[cm]
8.	CVDFAT02	Dieta niskotłuszczowa, niskocholesterolowa	1-tak/2-nie/0-nie wiem
wyjście	DIABETES	ZDIAGNOZOWANA CUKRZYCA CIĘŻARNYCH	1-tak/0-nie

Dokładną charakterystykę wyselekcjonowanych informacji, wykorzystanych w trakcie modelowania przedstawia Załącznik 2.

### 6.2.3. Przebieg prac

Przebieg prac obejmował modelowanie z wykorzystaniem danych pierwotnych oraz po zmianie charakterystyki części informacji wejściowych (zmiana charakteru informacji wejściowych).

Poniżej przedstawiono wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych pierwotnych.

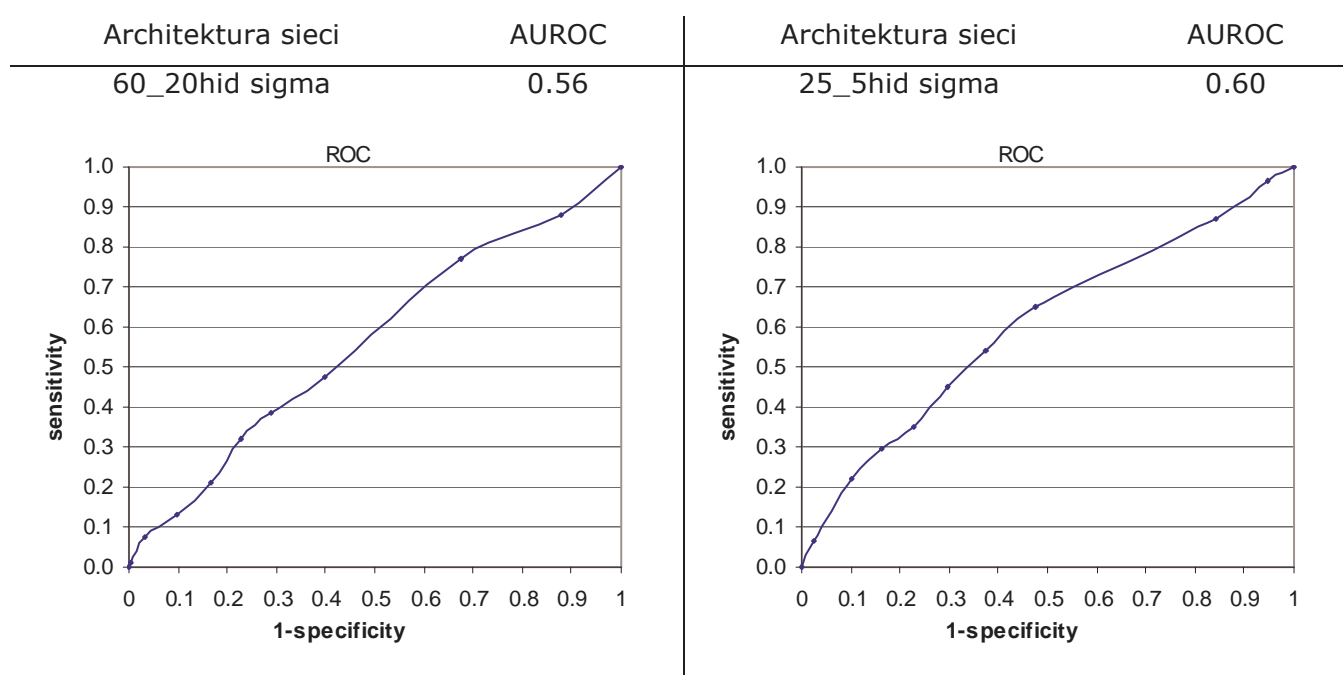
**Tabela 40** Najlepsze uzyskane wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych pierwotnych – dane balansowane ujemnie.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	20hid fsr	57	57.3	50.5	0.56
2	20_10hid tanh	54.9	55	51.6	0.56
3	30_5hid sigma	53.8	53.9	51.6	0.54
4	40hid sigma	53.6	53.8	49.5	0.54
5	50hid sigma	53.6	53.4	60.4	0.58

**Tabela 41** Najlepsze uzyskane wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych pierwotnych – dane balansowane dodatnio.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	60_20hid sigma	76.9	78.6	29.7	0.56
2	50_5hid sigma	74.8	76.6	26.4	0.55
3	18_8hid sigma	74.4	75.7	37.4	0.59
4	25_5hid sigma	73.9	75.2	38.5	0.60
5	20_10_2hid sigma	71.6	72.8	39.6	0.56

**Schemat 48** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka cukrzycy ciężarnych uzyskane na danych pierwotnych, balansowanych dodatnio.



Wyniki otrzymane na tym etapie pracy nie były satysfakcjonujące, dlatego podjęto próbę zmiany charakteru części danych wejściowych prezentowanych sieci neuronowej. Zmiana w danych wejściowych polegała na przekonwertowaniu dwóch zmiennych: *wiek* i *rasa*. W przypadku zmiennej *wiek* wartość ciągłą zastąpiono wartością przedziałową zgodnie z kluczem, który przedstawia Tabela 42.

**Tabela 42** Zakres zmian zmiennej *wiek* w problemie modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych przeformatowanych

Zakres wartości zmiennej w pliku pierwotnym	Wartość zmiennej przedziałowej
$\leq 25$	1
25-35	2
$\geq 35$	3

Dla zmiennej *rasa* zaproponowano wartość binarną - Tabela 43.

**Tabela 43** Zakres zmian zmiennej *rasa* w problemie modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych przeformatowanych

Zakres wartości zmiennej w pliku pierwotnym	Wartość zmiennej przedziałowej
kaukaska	1
inna	0

Najlepsze wyniki tej części doświadczeń przedstawiają odpowiednio Tabela 44 oraz Tabela 45 dla danych balansowanych ujemnie i dodatnio.

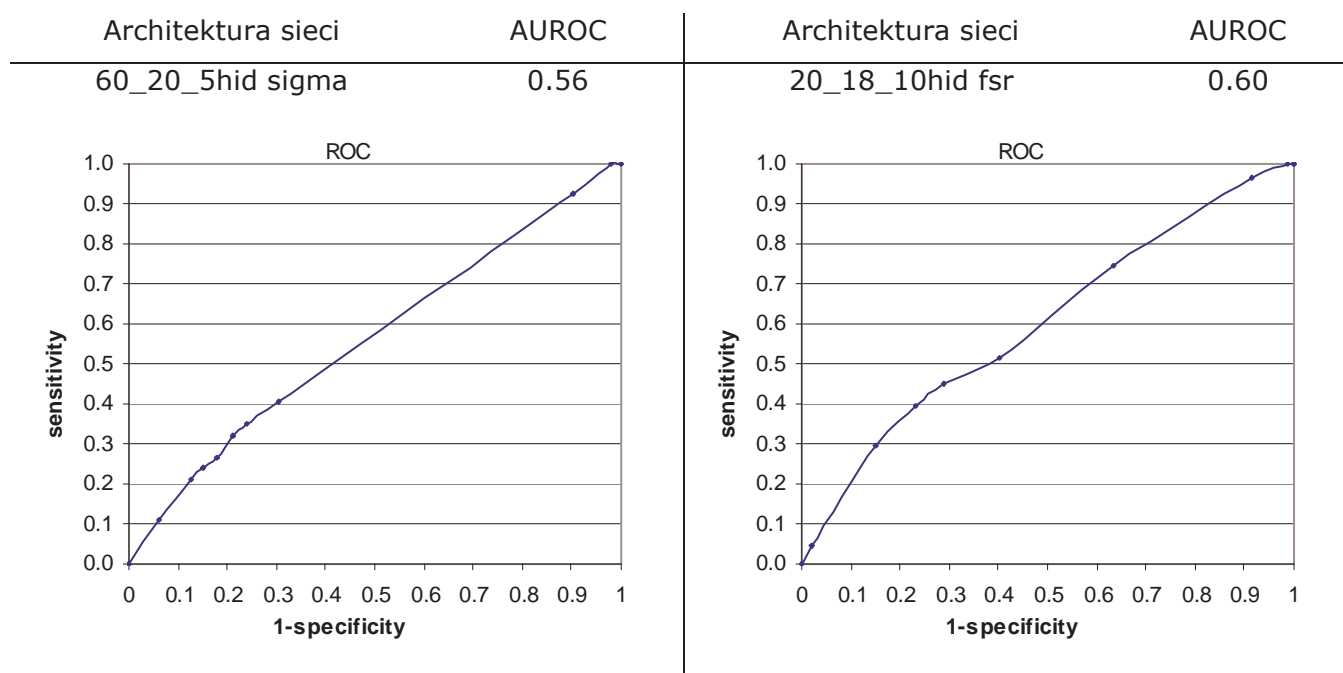
**Tabela 44** Najlepsze uzyskane wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych przekodowanych – dane balansowane ujemnie.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	20_10_2hid fsr	56.3	56.7	45.1	0.54
2	20_10_2hid sigma	55.5	55.7	52.7	0.54
3	7_5_3hid sigma	53.7	53.8	50.5	0.54
4	15_5hid sigma	53.1	53.1	51.6	0.54
5	30_10_5hid sigma	52.1	52.1	50.5	0.53

**Tabela 45** Najlepsze uzyskane wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych przekodowanych – dane balansowane dodatnio.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	60_20_5hid sigma	79.8	81.8	26.4	0.56
2	45_7_5hid sigma	78.0	79.6	35.2	0.58
3	18_8hid sigma	75.9	77.5	33.0	0.56
4	20_18_10hid fsr	71.1	72.1	45.1	0.60
5	20hid sigma	64.2	64.6	52.7	0.60

**Schemat 49** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka cukrzycy ciężarnych uzyskane na danych przekodowanych, balansowanych dodatnio.





## 6.3 Rezultaty poszczególnych etapów pracy – modelowanie ryzyka wystąpienia nadciśnienia tętniczego (BRFSS\_1999)

### 6.3.1 Opis problemu

Nadciśnienie jest najczęściej diagnozowaną chorobą na świecie [1] i jednocześnie najczęstszą przyczyną zgonów. Podstawowe czynniki ryzyka nadciśnienia obejmują [1,118,139]:

- wiek
- płeć
- parametr określany zbiorczo jako styl życia (dieta, nałogi, poziom aktywności fizycznej)
- czynniki genetyczne

Dodatkowymi czynnikami sprzyjającymi tak dramatycznej sytuacji są mała współpraca pacjentów w trakcie terapii (zjawisko non-compliance) oraz brak wyraźnych dla pacjenta objawów klinicznych, szczególnie we wczesnym stadium choroby [20,41]. Nawyk częstego pomiaru ciśnienia krwi, nawet u osób, u których nie zdiagnozowano do tej pory patologii ciśnienia tętniczego krwi nie jest wystarczająco szeroko rozpowszechniony. Wynikać to może po części z nowoczesnego stylu życia, nakładającego coraz większe wymagania dla organizmu i wiążącego się z tym braku czasu.

W związku z powyższym spostrzeżeniem wydaje się wskazanym próba konstrukcji powszechnie dostępnego, łatwego w obsłudze i wymagającego poświęcenia minimalnej ilości czasu pozwalającego zaklasyfikować pacjenta do grupy ryzyka. Efektem tego może okazać się skłonienie do wizyty w gabinecie lekarza i poddanie się klasycznej procedurze diagnostycznej. Proponowane narzędzie miałyby więc pełnić rolę prostego systemu skryningu nadciśnienia tętniczego. Postulowana powszechna dostępność może zostać osiągnięta z wykorzystaniem Internetu jako medium.

### 6.3.2 Przygotowanie danych

Zakres przygotowania danych obejmował dwie drogi wykonywanych doświadczeń. Pierwsza z nich obejmowała klasyczną metodykę polegającą na uczeniu sieci na pełnym zbiorze danych, po czym analizie wrażliwości zgodnie ze schematem, który przedstawiony został w rozdziale 2.1.8. Kolejny krok to redukcja rozmiaru wektora wejściowego do danych kluczowych wskazanych przez system. Druga procedura obejmuje ręczne przygotowanie kluczowych danych na podstawie analizy literatury.

Pierwotny zbiór BRFSS filtrowano pod kątem rodzaju odpowiedzi na jedno z pytań badania, które brzmiało: czy kiedykolwiek miałeś zdiagnozowane przez lekarza, pielęgniarkę lub innego specjalistę z zakresu ochrony zdrowia nadciśnienie tętnicze. Do dalszych prac zaklasyfikowano całość zbioru – 159 989 rekordów, z których każdy opisany był przez 217 zmiennych.

Jednym z założeń projektu była maksymalna prostota i łatwość użytkowania, więc redukcja wektora wejściowego jest konieczna nie tylko ze względu na wydajność obliczeniową. W pierwszym schemacie zredukowano wektor pierwotny do 38 (Tabela 46) oraz 8 (Tabela 47) parametrów kluczowych wykorzystując opisany wcześniej (Rozdział 2.1.8) moduł analizy wrażliwości.

**Tabela 46** Zmienne opisujące rekordy pacjentów włączonych do badania (nadciśnienie tętnicze, 38 wejść – po analizie wrażliwości)

L.p.	Kod zmiennej	Opis
1.	GENHLTH	Stan zdrowia
2.	HLTHPLAN	Ubezpieczenie zdrowotne
3.	MEDCOST	Okres czasu, w którym chory nie mógł pozwolić sobie na wizytę lekarską
4.	CHECKUP	Okres czasu od ostatniej rutynowej kontroli stanu zdrowia
5.	BPTAKE	Okres czasu od ostatniej rutynowej kontroli ciśnienia tętniczego krwi
6.	BLOODCHO	Czy poziom cholesterolu był kiedykolwiek kontrolowany

7.	TOLDHI	Zdiagnozowany podniesiony ponad normę poziom cholesterolu
8.	DIABETES	Zdiagnozowana cukrzyca
9.	RMVTEETH	Liczba usuniętych zębów
10.	SUNBURN	Oparzenia słoneczne w ciągu ostatnich 12 miesięcy
11.	DRINKANY	Ilość wypijanego alkoholu w trakcie kilku ostatnich miesięcy
12.	DRINKDRI	Jazda samochodem po spożyciu alkoholu
13.	AGE	Wiek
14.	EDUCA	Wykształcenie
15.	EMPLOY	Status pracowniczy
16.	WEIGHT	Waga
17.	SEX	Płeć
18.	FLUSHOT	Przypadki grypy w ciągu ostatnich 12 miesięcy
19.	PNEUMVAC	Czy kiedykolwiek podano szczepionkę przeciwko zapaleniu płuc
20.	SMKDETE2	Czas od ostatniego testu detektorów dymu tytoniowego
21.	RATECARE	Ocena satysfakcji z opieki zdrowotnej
22.	PRIMCARE	Opieka lekarska – samodzielny lekarz czy klinika
23.	FACILIT2	Miejsce do którego chodzisz najczęściej
24.	DISTCARE	Łatwość dostępu do usług medycznych
25.	CHNGDOCT	Zmiana lekarza prowadzącego
26.	PCSEATHB	Porady od lekarza na temat nawyków żywieniowych
27.	PCSEXER	Porady od lekarza na temat aktywności fizycznej
28.	CVDFATRK	Lekarz zalecił zastosowanie diety niskocholesterolowej
29.	CVDEXRSK	Lekarza zalecił zwiększenie aktywności fizycznej
30.	ASPUNSAF	Uczulenie lub nadwrażliwość na aspirynę

31.	WHYASPAN	Aspiryna stosowana jako lek przeciwbólowy
32.	LOSEWT	Próby obniżenia masy ciała
33.	MAINTAIN	Próby utrzymania obecnej masy ciała
34.	FEWCAL	Dieta niskokaloryczna lub/i niskotłuszczowa
35.	PHYACT	Zwiększona aktywność fizyczna
36.	WTDESIRE	Waga należna
37.	DRADVICE	Profesjonalna porada dotycząca żywienia
38.	USENOW	Aktualnie stosowane preparaty tytoniowe
Wyjście	BPHIGH	Zdiagnozowane nadciśnienie tętnicze

**Tabela 47** Zmienne opisujące rekordy pacjentów włączonych do badania (nadciśnienie tętnicze, 8 wejść – po analizie wrażliwości

L.p.	Kod zmiennej	Opis
1.	CHECKUP	Okres czasu od ostatniej rutynowej kontroli stanu zdrowia
2.	BPTAKE	Okres czasu od ostatniej rutynowej kontroli ciśnienia tętniczego krwi
3.	TOLDHI	Zdiagnozowany podniesiony ponad normę poziom cholesterolu
4.	RMVTEETH	Liczba usuniętych zębów
5.	DRINKANY	Ilość wypijanego alkoholu w trakcie kilku ostatnich miesięcy
6.	DRINKDRI	Prowadzenie samochodu po spożyciu alkoholu
7.	FLUSHOT	Przypadki grypy w ciągu ostatnich 12 miesięcy
8.	PNEUMVAC	Szczepienie p/grypie
Wyjście	BPHIGH	Zdiagnozowane nadciśnienie tętnicze

W drugim schemacie przeprowadzono analizę czynników ryzyka nadciśnienia oparta na piśmiennictwie i opiniach ekspertów. Wynikiem

prac był wybór sześciu najważniejszych z punktu widzenia ryzyka nadciśnienia zmiennych z dostępnego spektrum.

**Tabela 48** Zmienne opisujące rekordy pacjentów włączonych do badania (nadciśnienie tętnicze, 6 wejść - ręczne)

L.p.	Kod zmiennej	Opis
1.	TOLDHI	Zdiagnozowany podniesiony ponad normę poziom cholesterolu
2.	SMOKENUM	Liczba wypalanych dziennie papierosów
3.	AGE	Wiek
4.	WEIGHT	Waga
5.	HEIGHT	Wzrost
6.	SEX	Płeć
Wyjście	BPHIGH	Zdiagnozowane nadciśnienie tętnicze

Podobnie jak w problemie dotyczącym predykcji ryzyka cukrzycy ciężarnych, ze względu na nierówny rozkład liczebności przypadków konieczne było balansowanie klas – przeprowadzane analogicznie jak opisano w rozdziale 6.2.2.

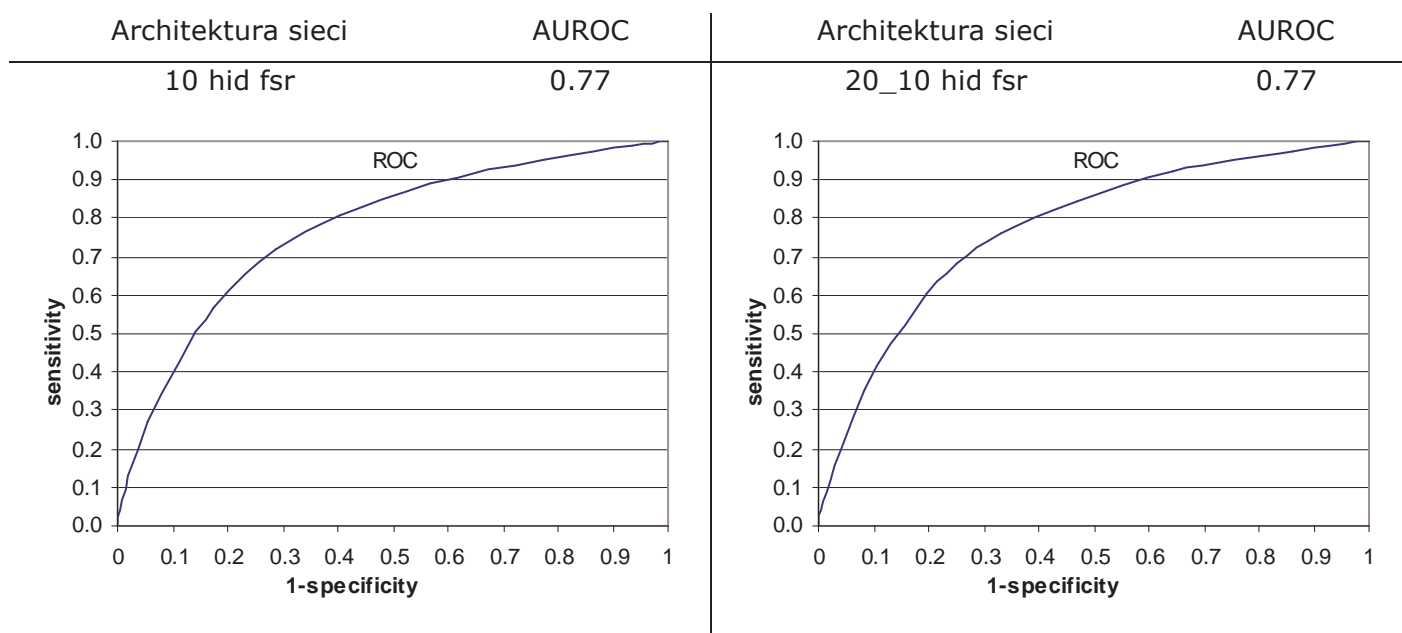
### 6.3.3. Przebieg prac

Poniższe tabele (Tabela 49, Tabela 50) przedstawiają najlepsze wyniki uzyskane z wykorzystaniem danych uzyskane po analizie wrażliwości, opartych na 38 zmiennych wejściowych odpowiednio balansowanych ujemnie i dodatnio. Dodatkowo Schematy 47 oraz 48 zawierają wykresy ROC dla najlepszych uzyskanych modeli.

**Tabela 49** Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych pierwotnych (38 wejść) – dane balansowane ujemnie.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	10 hid fsr	75.45	80.4	61.0	0.77
2	20_10 hid fsr	73.89	76.4	66.7	0.77
3	10 hid tanh	73.73	76.1	66.8	0.77
4	5_3 hid fsr	73.18	74.7	68.8	0.77
5	7_5_3 hid fsr	72.66	73.8	69.4	0.77

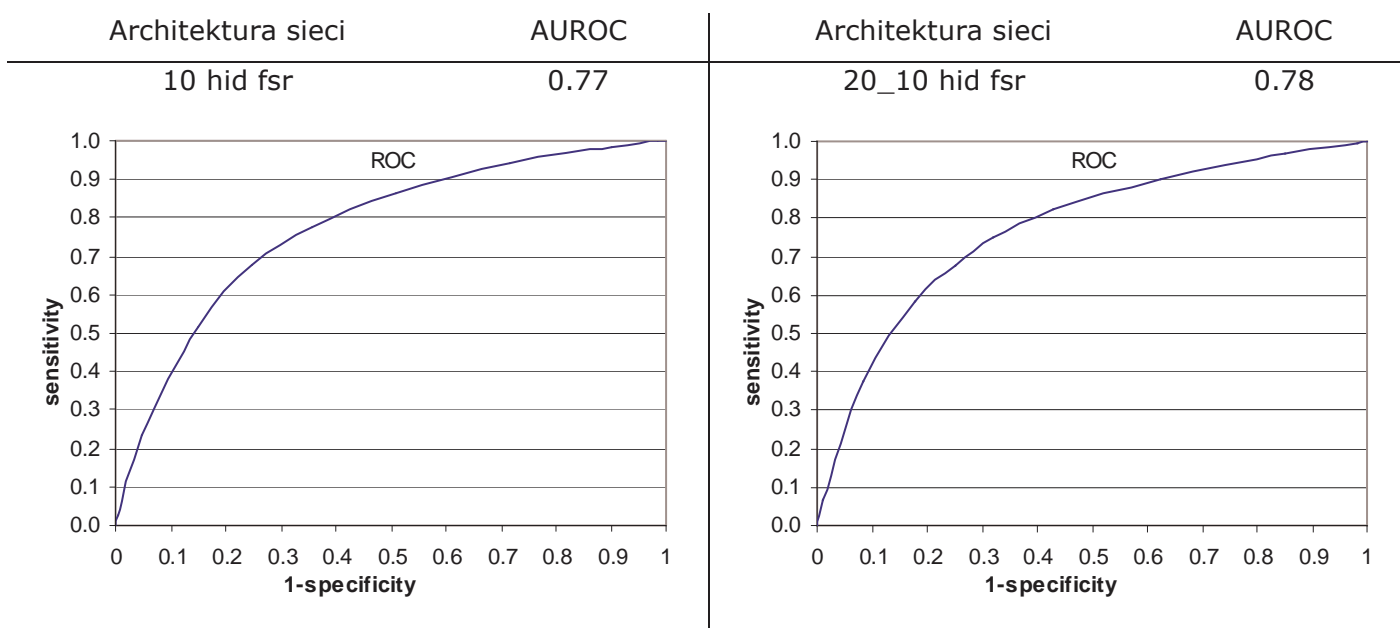
**Schemat 50** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych pierwotnych (38 wejść), balansowanych ujemnie.



**Tabela 50** Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych pierwotnych – dane balansowane dodatnio.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	10 hid fsr	75.13	79.7	61.8	0.77
2	20_10 hid fsr	73.49	75.6	67.5	0.78
3	5_3 hid fsr	73.42	75.3	68.0	0.77
4	30_20_10 hid fsr	73.09	74.6	68.7	0.77
5	10 hid tanh	73.06	74.7	68.3	0.77

**Schemat 51** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych pierwotnych (38 wejść), balansowanych dodatnio.

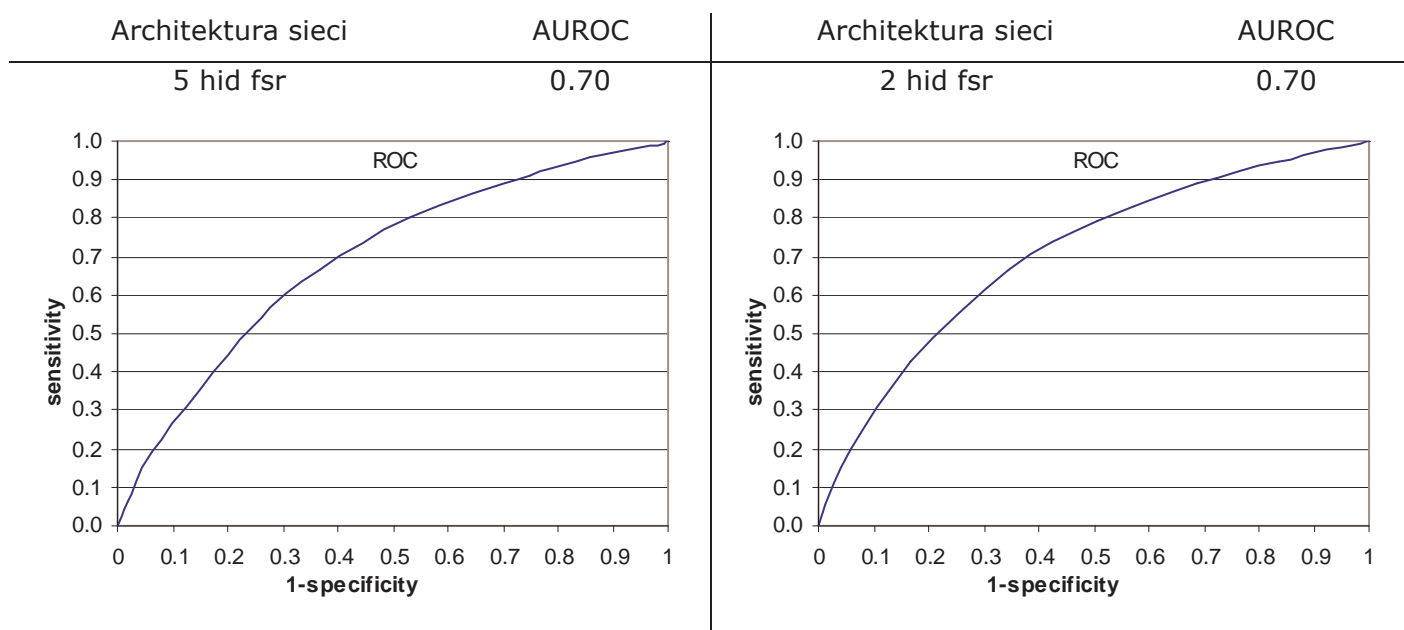


Kolejny krok to modelowanie ryzyka wystąpienia nadciśnienia tętniczego w oparciu o dane uzyskane również po analizie wrażliwości jednak dla wektora wejściowego zredukowanego do 8 wejść.

**Tabela 51** Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych pierwotnych (8 wejść) – dane balansowane ujemnie.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	5 hid fsr	71.01	78.23	49.82	0.70
2	2 hid fsr	70.71	77.56	50.6	0.70
3	10 hid fsr	70.18	76.2	52.52	0.70
4	2 hid tanh	69.47	74.56	54.61	0.70
5	3_2 hid tanh	68.53	70.4	63.07	0.71

**Schemat 52** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych pierwotnych (8 wejść), balansowanych ujemnie.

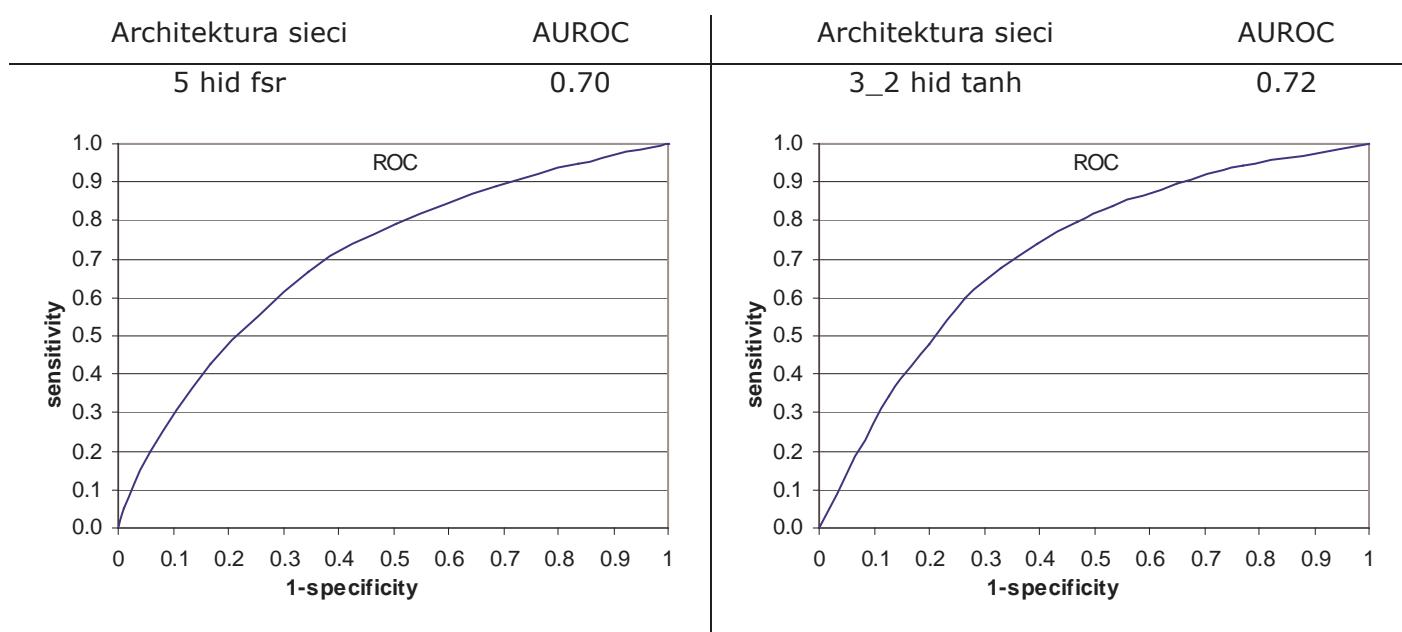




**Tabela 52** Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych pierwotnych (8 wejść) – dane balansowane dodatnio.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	5 hid fsr	70.66	77.46	50.73	0.70
2	2 hid tanh	70.12	76.26	52.07	0.70
3	2 hid fsr	70.07	76.06	52.44	0.70
4	10 hid fsr	69.93	75.58	53.4	0.70
5	3_2 hid tanh	69.81	72.72	61.29	0.72

**Tabela 53** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych pierwotnych (8 wejść), balansowanych dodatnio.



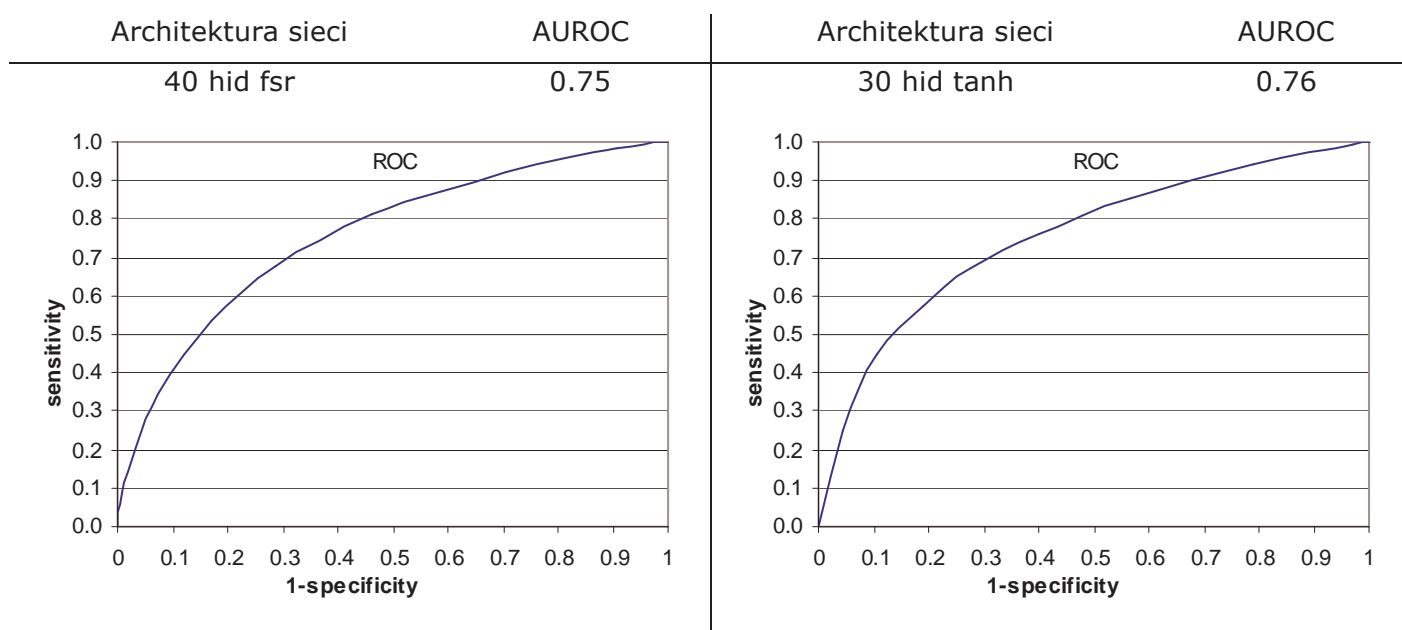
W celu porównania wyników uzyskiwanych na drodze automatycznej redukcji wektora wejściowego oraz po jego ręcznej, arbitralnej obróbce wykonano doświadczenie polegające na przygotowaniu danych o możliwie małym wektorze wejściowym. Efektem prac było przygotowanie par plików testowy-uczący bazujących na 6 zmiennych sterujących, otrzymanych w wyniku analizy danych literaturowych. Tabela 54 oraz Tabela 56

przedstawiają najlepsze uzyskane wyniki dla odpowiednio zbioru balansowanego ujemnie i dodatnio, natomiast Tabela 55 i Tabela 57 wykresy ROC najlepszych sieci.

**Tabela 54** Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych uzyskanych w oparciu o analizę literatury (6 wejść) – dane balansowane ujemnie.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	40hid fsr_6in_all	72.98	82.42	49.93	0.75
2	30hid tanh_6in_all	70.29	66.37	71.9	0.76
3	25hid fsr_6in_all	69.94	67.18	71.11	0.76
4	25hid tanh_6in_all	69.91	66.29	71.39	0.76
5	15hid tanh_30in_all	68.93	71.32	67.95	0.76

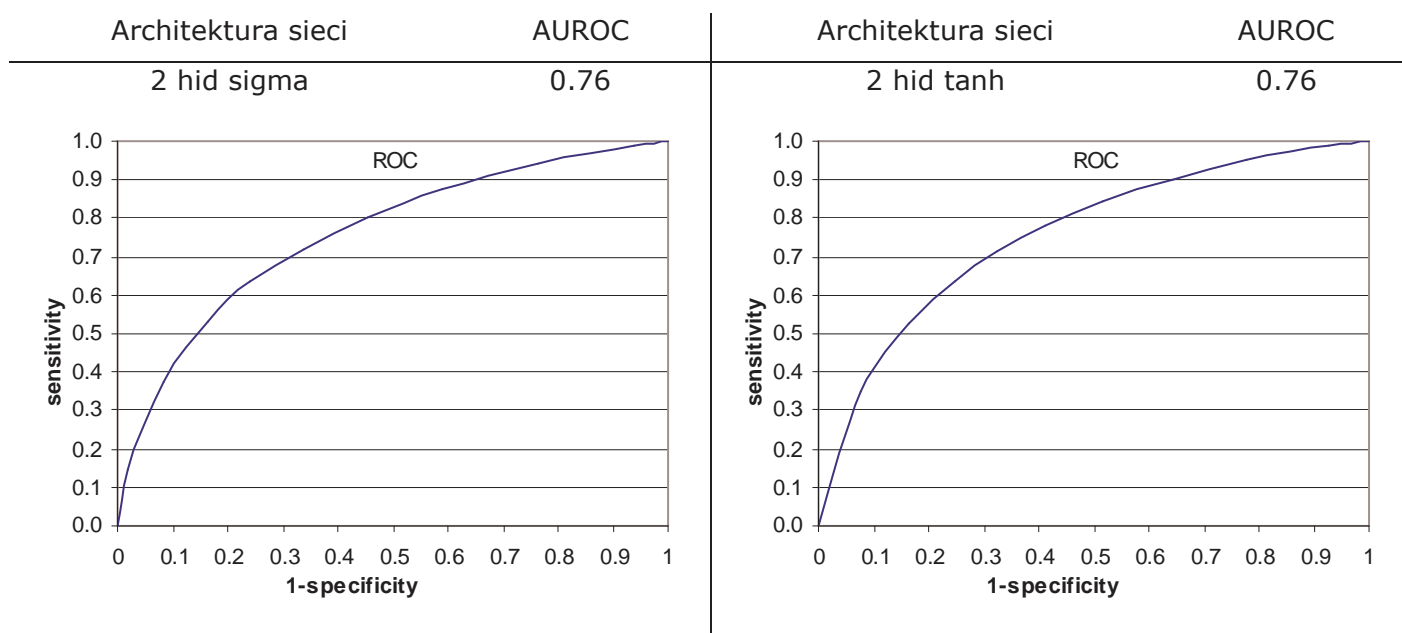
**Tabela 55** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych uzyskanych w oparciu o analizę literatury (6 wejść), balansowanych ujemnie.



**Tabela 56** Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych uzyskanych w oparciu o analizę literatury (6 wejść) – dane balansowane dodatnio.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	2hid sigma_6in_all	72.91	58.39	77.82	0.76
2	2hid tanh_6in_all	71.27	62.99	74.66	0.76
3	20_10hid tanh_6in_all	71.14	63.59	74.25	0.76
4	10hid tanh_6in_all	71.11	63.3	74.31	0.76
5	10hid fsr_6in_all	71.02	63.92	73.93	0.76

**Tabela 57** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych uzyskanych w oparciu o analizę literatury (6 wejść), balansowanych dodatnio.



## 6.4 Rezultaty poszczególnych etapów pracy – modelowanie ryzyka wystąpienia ostrego incydentu wieńcowego (NHEFS I)

### 6.4.1. Opis problemu

Choroby układu krążenia są dominującą grupą schorzeń wśród populacji krajów rozwiniętych. Zmiany stylu życia wynikające z

postępującej industrializacji powodują, że tego typu schorzenia są najczęstszą przyczyną śmierci.

Definicja ataku serca lub zawału serca są dyskutowane [34]. Najczęściej stosuje się pojęcie ostry incydent lub ostry zespół wieńcowy, co oznacza martwicę kardiomiocytów, której przyczyną jest długotrwałe niedokrwienie. Diagnostyka polega zarówno na ocenie objawów klinicznych jak i poziomu markerów biochemicznych. Podstawowe czynniki ryzyka ostrego incydentu wieńcowego obejmują dwie podstawowe grupy określane jako modyfikowalne i niemodyfikowalne [66]. Do pierwszej zalicza się przede wszystkim:

- nadciśnienie tętnicze
- hipercholesterolemia/hipertrójglicerydemia
- palenie tytoniu
- poziom aktywności fizycznej
- otyłość, dieta
- cukrzyca

Drugą grupę stanowią:

- wiek
- czynniki genetyczne
- płeć
- rasa

W związku z obecnością licznych czynników podlegających modyfikacji oraz wagi samego problemu, prewencja i wczesna selekcja osób obarczonych ryzykiem zaistnienia ostrego incydentu wieńcowego jest obecnie wyraźnie podkreślana. Oprócz istnienia licznych programów o charakterze promocji zdrowia, podkreśla się rolę konstrukcji możliwie prostych modeli predykcyjnych, mogących stanowić narzędzie przesiewowe. Przykładem takiego systemu może być dostępny on-line, a oparty o badanie Framingham Heart Study kalkulator internetowy badający 10-cio letnie ryzyko ataku serca [28]. Zaletą przedstawionego narzędzia jest jego dostępność oraz duża łatwość używania. Jako

ewentualne minusy wskazać można konieczność znajomości wartości ciśnienia oraz poziomu cholesterolu całkowitego i frakcji HDL.

Wymienione niedogodności mogą stanowić punkt wyjścia dla próby zastosowania sztucznych sieci neuronowych jako narzędzi komplementarnych w stosunku do modelu liniowego. Konstrukcja podobnego systemu musi się więc opierać na jeszcze łatwiej dostępnych dla przeciętnego pacjenta parametrach – nie tracąc przy tym skuteczności predykcyjnej.

#### 6.4.2. Przygotowanie danych

Wykorzystany w trakcie prac zbiór NHANES filtrowano pod kątem obecności odpowiedzi przy pytaniu dotyczącym ataku serca. Do badania włączono te przypadki, w których udało się zidentyfikować istnienie jednoznacznie pozytywnej lub negatywnej odpowiedzi. Wyłączono przypadki zawierające odpowiedź niejednoznaczną (nie wiem lub sprzeczną z pytaniami kontrolnymi). W wyniku powyższej procedury wyselekcjonowano zbiór zawierający 740 rekordów. Pierwotnie każdy z respondentów opisany był przez 1287 zmiennych. W opisanym powyżej zbiorze konieczna była redukcja wymiarów ze względu na istnienie informacji pustych (ta sama wartość dla wszystkich wyselekcjonowanych 740 rekordów). Ostatecznie rozmiar wektora pierwotnego wynosił 858 zmiennych sterujących.

Wyraźna dysproporcja między ilością rekordów i opisujących ich zmiennych była przyczyną kolejnej redukcji rozmiaru wektora wejściowego. Podobnie jak w doświadczeniu opisanym poprzednio przetestowano dwie możliwe ścieżki – analizę wrażliwości i ręczne przygotowanie danych. Analiza wrażliwości została przeprowadzona zgodnie z metodyką, która przedstawia rozdział 2.1.8. Ręcznie wyselekcjonowane dane odpowiadały podstawowym czynnikom ryzyka opisanym w piśmiennictwie. W pierwszym schemacie zredukowano wektor pierwotny do 21 wejść (Tabela 58) natomiast w drugim do 8 wejść (Tabela 59).

**Tabela 58** Zmienne opisujące rekordy pacjentów włączonych do badania (atak serca, 28 wejść – po analizie wrażliwości)

L.p.	Kod zmiennej	Opis
1.	LNURSING	Rezydent domu opieki
2.	LWORKHST	Respondent czynny zawodowo
3.	LHICHOL	Podniesiony ponad normę poziom cholesterolu
4.	LBCANCYR	Rok diagnozy raka piersi
5.	LBROKHIP	Złamanie kości udowej
6.	LVASECT	Wazektomia
7.	LCALCIUM	Stosowanie preparatów wapnia
8.	LONEOVAR	Usunięcie jajnika
9.	ORALDB	Stosowanie doustnych leków hipoglikemizujących
10.	HOSPFLU	Hospitalizacja z powodu zapalenia płuc.
11.	HERNIA	Zdiagnozowana w przeszłości przepuklina
12.	CATRSURG	Zabieg chirurgiczny usunięcia katarakty
13.	PREVYRHT	Stan zdrowia 12 miesięcy przed przyjęciem do domu opieki
14.	SUBJCOND	Przyczyna zgonu
15.	OFTNWINE	Spożycie wina w ostatnich 12 miesiącach
16.	MEDICARE	Ubezpieczenie zdrowotne
17.	SAVINGS	Dochody z oprocentowanych lokat
18.	DIVIDEND	Dochody z tytułu dywidend
19.	CURRSX1	Płeć członka gospodarstwa domowego
20.	HCFSR04	Pobyty w szpitalu z powodu innych poza incydem wieńcowym niedomogów serca
21.	HCFSR08	Pobyty w szpitalu z powodu zaburzeń krążenia mózgowego
22.	HCFSR09	Pobyty w szpitalu z powodu cukrzycy

23.	HCFSR11	Pobył w szpitalu z powodu chorób nowotworowych
24.	HCFSR23	Pobył w szpitalu z powodu bólu/wycieńczenia
25.	HCFSR25	Pobył w szpitalu z powodu kłopotów z układem mięśniowo szkieletowym
26.	HCFSR30	Pobył w szpitalu z powodu odwodnienia lub niedożywienia
27.	HCFSR35	Pobył w szpitalu
28.	HCFSR37	Pobył w szpitalu z powodu upadku
Wyjście	HEARTATK	Ostry incydent wieńcowy

**Tabela 59** Zmienne opisujące rekordy pacjentów włączonych do badania (atak serca, 8 wejść - ręczne).

L.p.	Kod zmiennej	Opis
1.	LSSEX	Płeć
2.	LHICHOL	Podniesiony ponad normę poziom cholesterolu
3.	AGE92	Wiek
4.	SUGRDIAB	Kiedykolwiek zdiagnozowana cukrzyca
5.	HYPER	Kiedykolwiek zdiagnozowane nadciśnienie tętnicze krwi
6.	SMOKENOW	Palenie papierosów
7.	EXERNUM2	Aktywność fizyczna (miesięczna/tygodniowa)
8.	WEIGHNOW	Waga
Wyjście	HEARTATK	Ostry incydent wieńcowy

Ze względu na nierówny rozkład liczebności przypadków – testowano zastosowane wcześniej balansowanie klas.

#### 6.4.3. Przebieg prac

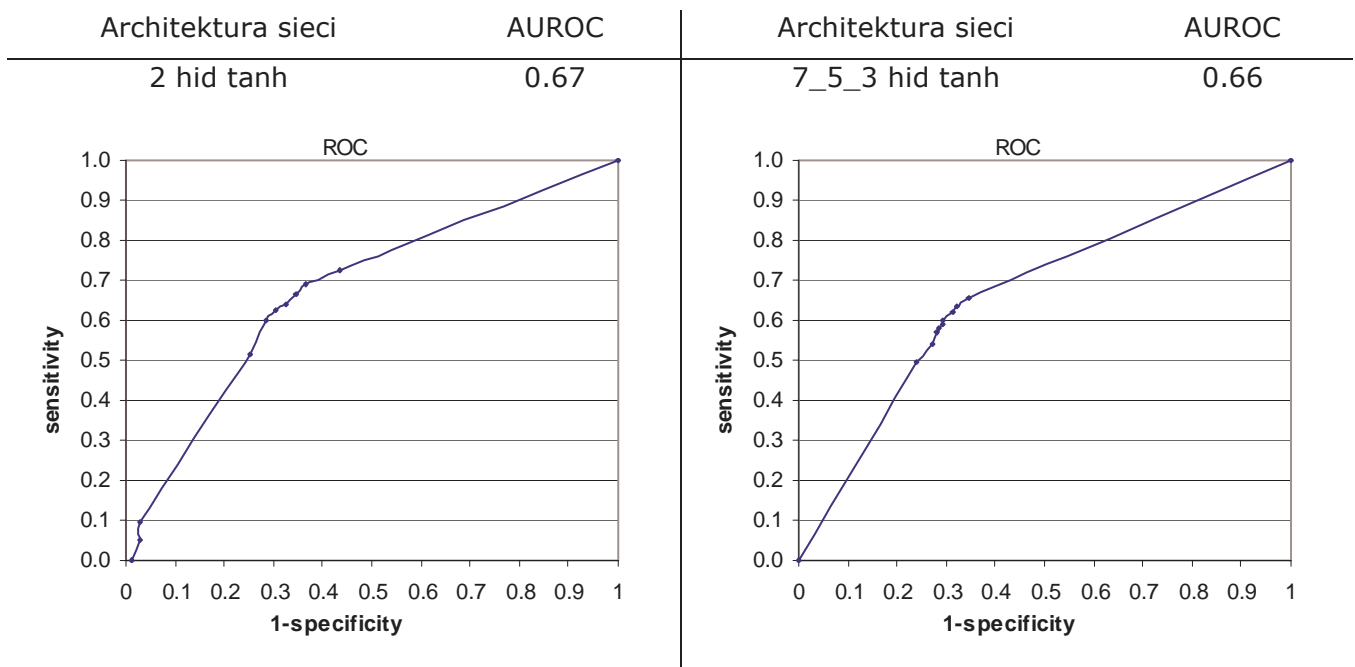
Poniższe tabele (Tabela 60, Tabela 61) przedstawiają najlepsze wyniki uzyskane z wykorzystaniem danych uzyskanych po analizie wrażliwości, opartych na 28 zmiennych wejściowych odpowiednio

balansowanych ujemnie i dodatnio. Dodatkowo oraz zawierają wykresy ROC dla najlepszych uzyskanych modeli (Schemat 50, Schemat 51).

**Tabela 60** Najlepsze uzyskane wyniki modelowania ryzyka ostrego incydentu wieńcowego z wykorzystaniem danych pierwotnych (28 wejść) – dane balansowane ujemnie.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	2 tanh_out_HEARTATK_only_pos_28in_bal-	65.00	66.67	64.59	0.67
2	7_5_3 tanh_out_HEARTATK_only_pos_28in_bal-	61.49	70.75	59.19	0.66
3	25 tanh_out_HEARTATK_only_pos_28in_bal-	61.49	65.99	60.37	0.65
4	5_3 tanh_out_HEARTATK_only_pos_28in_bal-	60.68	70.75	58.18	0.66
5	15_7_3 tanh_out_HEARTATK_only_pos_28in_bal-	60.68	63.95	59.87	0.64

**Schemat 53** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka ostrego incydentu wieńcowego uzyskane na danych pierwotnych (28 wejść), balansowanych ujemnie.

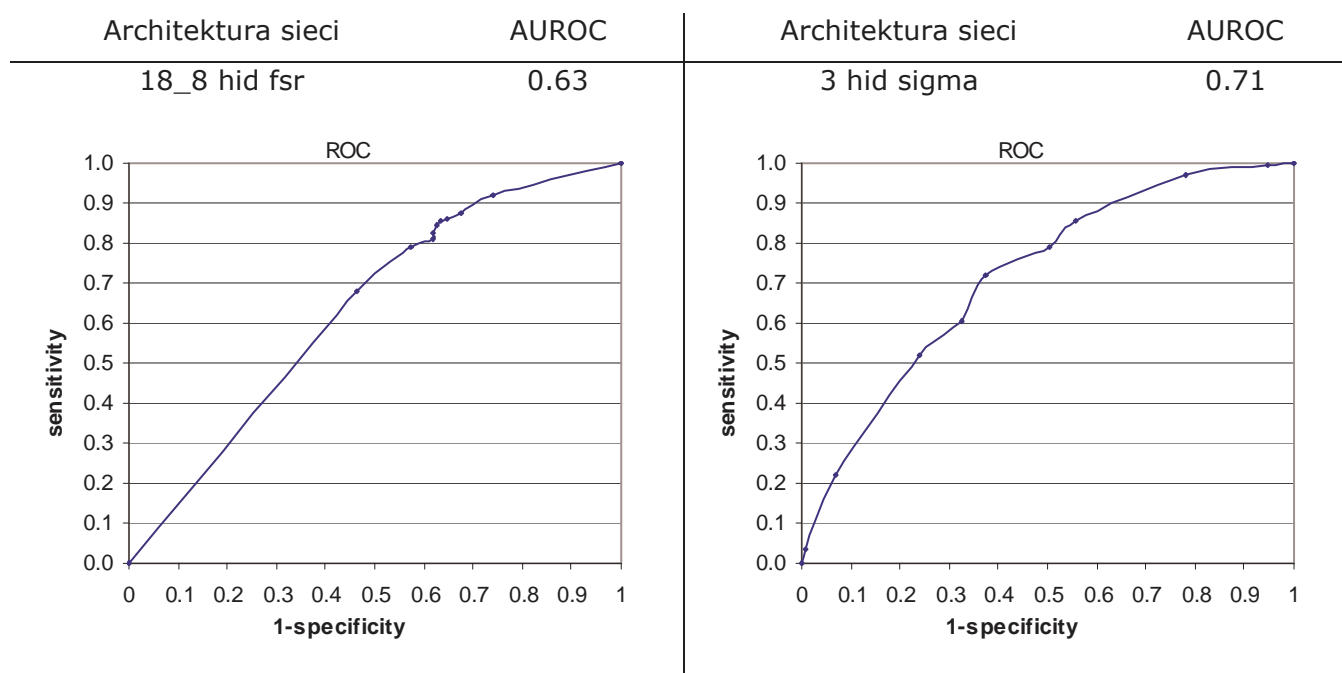




**Tabela 61** Najlepsze uzyskane wyniki modelowania ryzyka ostrego incydentu wieńcowego z wykorzystaniem danych pierwotnych (28 wejść) – dane balansowane dodatnio.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	18_8 fsr_out_HEARTATK_only_pos_28in_bal+	75.14	37.41	84.49	0.63
2	40_15 sigma_out_HEARTATK_only_pos_28in_bal+	74.73	42.86	82.63	0.64
3	5_3 tanh_out_HEARTATK_only_pos_28in_bal+	73.51	57.14	77.57	0.68
4	2 sigma_out_HEARTATK_only_pos_28in_bal+	71.35	55.10	75.38	0.72
5	3 sigma_out_HEARTATK_only_pos_28in_bal+	70.14	62.59	72.01	0.71

**Schemat 54** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka ostrego incydentu wieńcowego uzyskane na danych pierwotnych (28 wejść), balansowanych dodatnio.



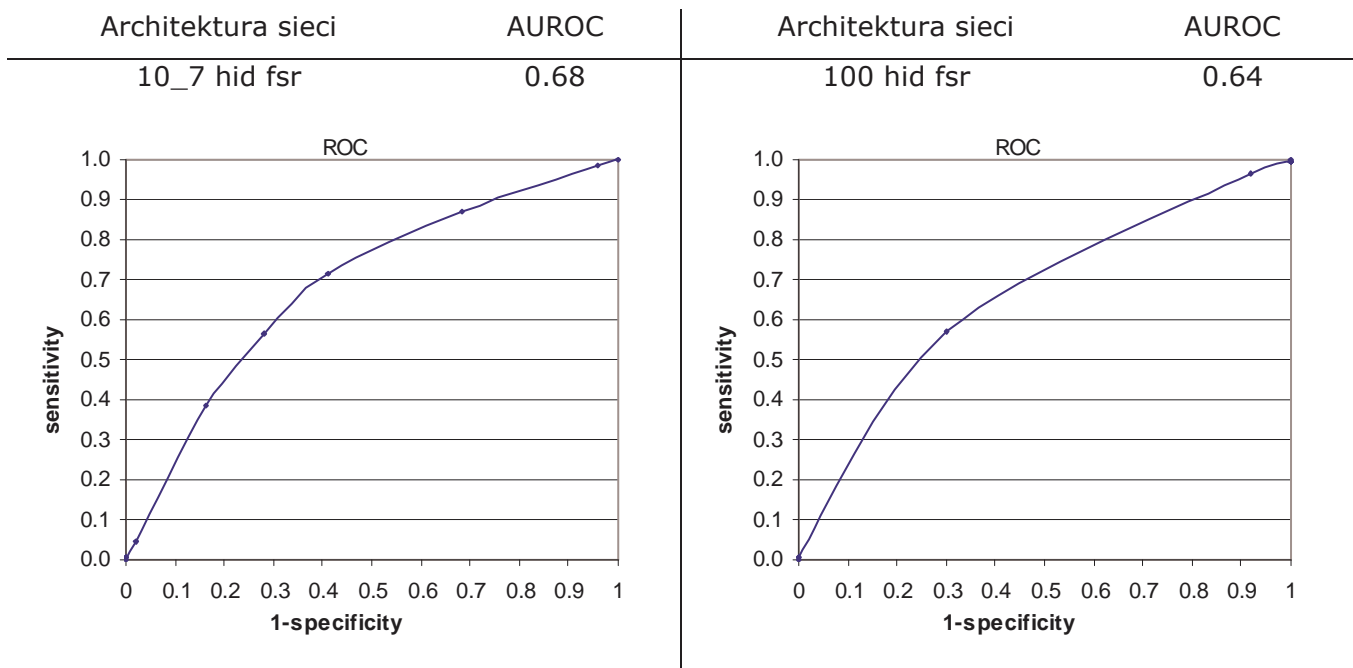
W celu porównania wyników modeli zbudowanych w oparciu o analizę wrażliwości przygotowano zbiór danych uzyskanych w wyniku analizy danych literaturowych. Poniższe tabele (Tabela 62, Tabela 63) przedstawiają najlepsze wyniki uzyskane z wykorzystaniem danych opartych na 8 zmiennych wejściowych odpowiednio balansowanych

ujemnie i dodatnio. Dodatkowo oraz zawierają wykresy ROC dla najlepszych uzyskanych modeli (Schemat 52, Schemat 53).

**Tabela 62** Najlepsze uzyskane wyniki modelowania ryzyka ostrego incydentu wieńcowego z wykorzystaniem danych pierwotnych (8 wejść) – dane balansowane ujemnie.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	10_7_fsr_30in_all	62.4	67.8	61.1	0.68
2	100_fsr_30in_all	62.4	65.8	61.6	0.64
3	20_fsr_30in_all	62.4	67.8	61.1	0.65
4	100_tanh_30in_all2	62.2	73.2	59.4	0.68
5	50_fsr_30in_all	62.2	64.4	61.6	0.64

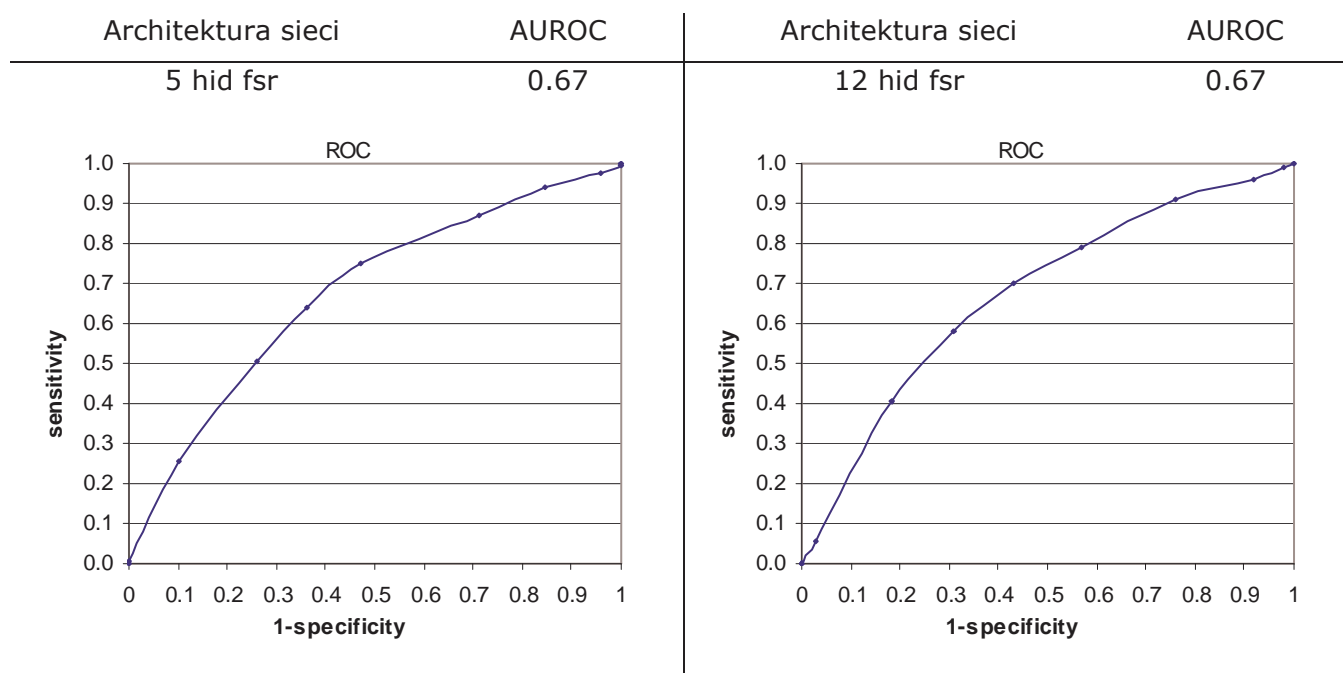
**Schemat 55** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka ostrego incydentu wieńcowego uzyskane na danych pierwotnych (8 wejść), balansowanych ujemnie.



**Tabela 63** Najlepsze uzyskane wyniki modelowania ryzyka ostrego incydentu wieńcowego z wykorzystaniem danych pierwotnych (28 wejść) – dane balansowane dodatnio.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	5_fsr_30in_all	67.7	57.7	70.2	0.67
2	12_fsr_30in_all	67.4	57	70.1	0.67
3	12_tanh_30in_all	67.4	56.4	70.2	0.67
4	10_fsr_30in_all	67	61.7	68.4	0.68
5	100_tanh_30in_all	66.8	55	69.7	0.68

**Schemat 56** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka ostrego incydentu wieńcowego uzyskane na danych pierwotnych (28 wejść), balansowanych dodatnio.



## 6.5 Rezultaty poszczególnych etapów pracy – modelowanie ryzyka wystąpienia choroby nowotworowej (NHEFS I)

### 6.5.1. Opis problemu

Obok chorób układu krążenia choroby nowotworowe stanowią jedną z najczęstszych przyczyn zgonów na świecie [67]. Samo przedstawienie zagadnień i sposobów postępowania z tym zjawiskiem zajmuje setki stron międzynarodowych opracowań. Rozwój metod terapeutycznych umożliwia wydłużenie przeżyci lub nawet wyleczenie, niemniej jednak podkreślana jest przede wszystkim rola prewencji, profilaktyki oraz rozwój metod wczesnej diagnostyki, bez względu na typ nowotworu [106]. Wyniki badań epidemiologicznych oraz konstruowane na ich podstawie wytyczne stawiają powyższe czynniki na najbardziej eksponowanych miejscach [4].

Wczesna diagnostyka jest na tyle ważna z prognostycznego punktu widzenia, że w celu zwiększenia dostępności konstruowane są narzędzie internetowe wspomagające diagnostykę lub informujące o ryzyku danego nowotworu. Przykładem może być model skonstruowany na podstawie danych epidemiologicznych National Cancer Institute (NIC) [63] - kalkulator ryzyka nowotworu piersi u kobiet [58] lub interaktywna strona przygotowana przez Harvard School of Public Health [68].

Większość podobnych modeli jest zbudowana z wykorzystaniem klasycznej drogi modelowania z wykorzystaniem technik statystycznych. Wydaje się, iż dzięki swoim właściwościom metody oparte o techniki sztucznej inteligencji, mogłyby pozwolić na nowe spojrzenie na opisywane problemy.

### 6.5.2. Przygotowanie danych

Zbiór NHANES filtrowany był pod kątem obecności odpowiedzi respondenta na pytanie dotyczące choroby nowotworowej, bez względu na jej typ i rodzaj. Do badania włączone te przypadki, w których zidentyfikowano pozytywną lub negatywną odpowiedź. Kryterium wyłączenia była odpowiedź niejednoznaczna (nie wiem) lub sprzeczna

(istniała możliwość wychwycenia sprzeczności dzięki obecności pytań kontrolnych).

W wyniku powyższej procedury wyselekcjonowano zbiór zawierający 9281 rekordów. Pierwotnie każdy z respondentów opisany był przez 878 zmiennych. W opisanym powyżej zbiorze konieczna była redukcja wymiarów ze względu na istnienie informacji pustych (ta sama wartość dla wszystkich wyselekcjonowanych rekordów). Ostatecznie rozmiar wektora pierwotnego wynosił 514 zmiennych sterujących. Ze względu na szeroki zakres modelowanego problemu zrezygnowano z ręcznego przygotowania zbiorów. Przeprowadzona została analiza wrażliwości wejść (Rozdział 2.1.8) i w kolejnych krokach zredukowano rozmiar wektora wejściowego testując skuteczność zbudowanych na ich podstawie modeli. Dokładny opis poszczególnych zbiorów danych i zmiennych sterujących w kolejnych krokach (101 wejść, 18 wejść) przedstawia Załącznik 3.

Ze względu na dysproporcje w rozkładzie zmiennej wyjściowej, we wszystkich przypadkach testowano balansowanie ujemne i dodatnie.

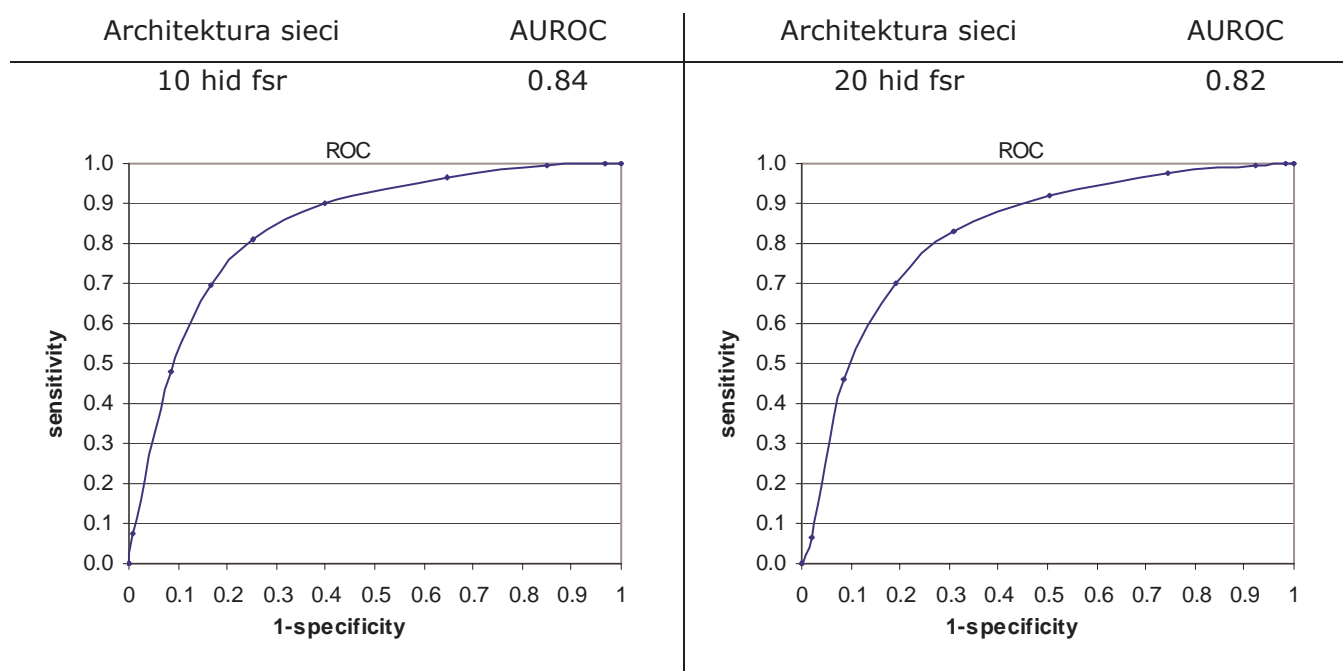
### 6.5.3. Przebieg prac

Poniższe tabele przedstawiają najlepsze wyniki uzyskane z wykorzystaniem danych opisanych przez 101 zmiennych sterujących uzyskanych w wyniku analizy wrażliwości, odpowiednio dla danych balansowanych ujemnie (Tabela 64, Schemat 54) i dodatnio (Tabela 65, Schemat 55).

**Tabela 64** Najlepsze uzyskane wyniki modelowania ryzyka wystąpienia choroby nowotworowej z wykorzystaniem danych pierwotnych (101 wejść) – dane balansowane ujemnie.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	10 hid fsr_out_CANCRSRT_101_in_bin	79.7	75.2	80	0.84
2	20 hid fsr_out_CANCRSRT_101_in_bin	79.4	72.6	79.8	0.82
3	2 hid fuzzy_out_CANCRSRT_101_in_bin_MSLIDE	78.3	73.1	78.6	0.82
4	2 hid tanh_out_CANCRSRT_101_in_bin	78.2	73.6	78.5	0.82
5	20 hid fuzzy_out_CANCRSRT_101_in_bin_MSLIDE	78.2	73.3	78.5	0.82

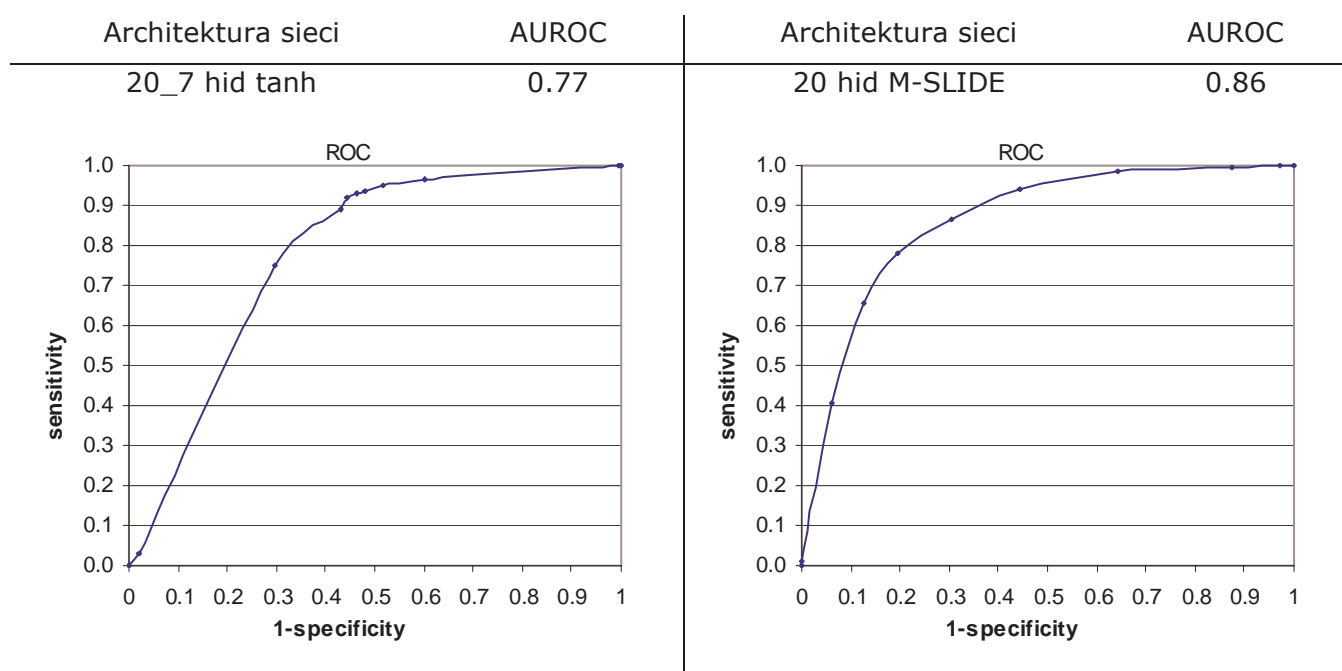
**Schemat 57** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka wystąpienia choroby nowotworowej uzyskane na danych pierwotnych (101 wejść), balansowanych ujemnie.



**Tabela 65** Najlepsze uzyskane wyniki modelowania ryzyka wystąpienia choroby nowotworowej z wykorzystaniem danych pierwotnych (101 wejść) – dane balansowane dodatnio.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	20_7 hid tanh_out_CANCRSRT_101_in_bin	90.7	53.6	92.8	0.77
2	15_7 hid fsr_out_CANCRSRT_101_in_bin	88.5	59.5	90.2	0.81
3	20 hid fuzzy_out_CANCRSRT_101_in_bin_MSLIDE	82.6	74.5	83.1	0.86
4	10 hid fsr_out_CANCRSRT_101_in_bin	82.5	74.7	82.9	0.86
5	3_2 hid tanh_out_CANCRSRT_101_in_bin	82.3	74.5	82.7	0.86

**Schemat 58** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka wystąpienia choroby nowotworowej uzyskane na danych pierwotnych (101 wejść), balansowanych dodatnio.

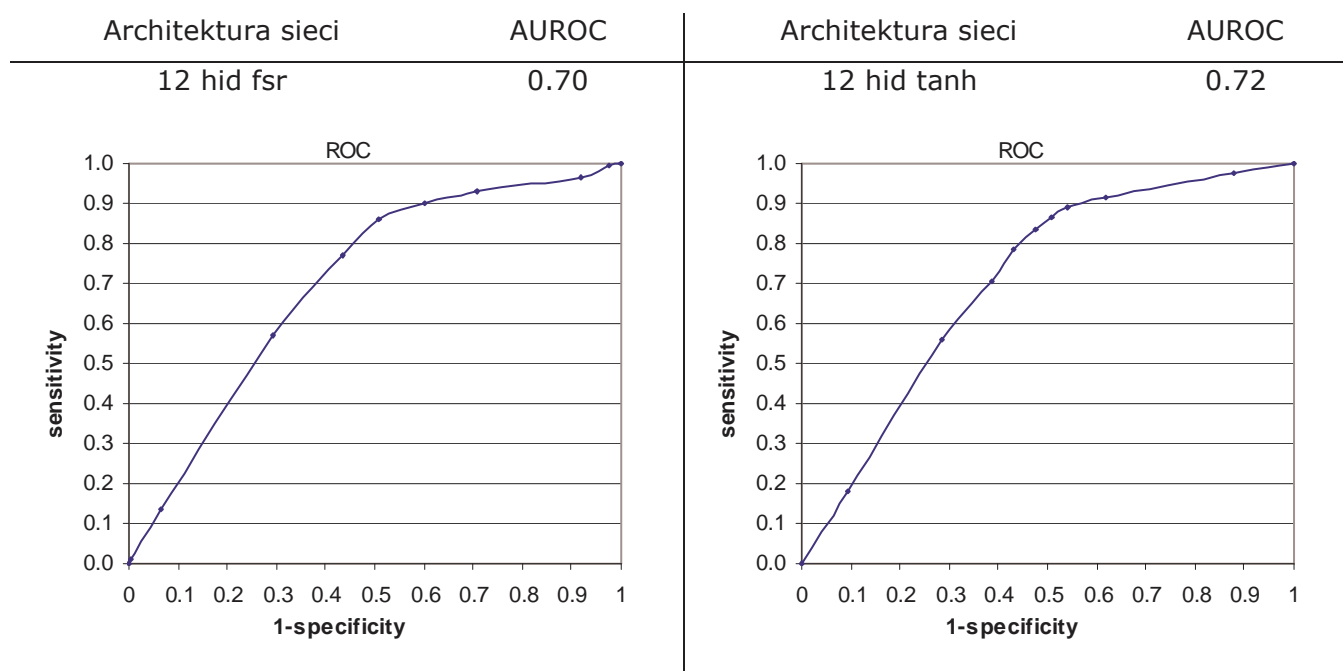


Kolejny etap to dalsza redukcja wektora wejściowego. Poniższe tabele przedstawiają najlepsze wyniki uzyskane z wykorzystaniem danych opisanych przez 18 zmiennych sterujących odpowiednio dla danych balansowanych ujemnie (Tabela 66, Schemat 56) i dodatnio (Tabela 67, Schemat 57).

**Tabela 66** Najlepsze uzyskane wyniki modelowania ryzyka wystąpienia choroby nowotworowej z wykorzystaniem danych pierwotnych (18 wejść) – dane balansowane ujemnie.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	12_fsr_CANCRSRT_18in_all	82.7	50.3	84.5	0.7
2	2_fsr_CANCRSRT_18in_all	82.6	50.1	84.5	0.7
3	50_tanh_CANCRSRT_18in_all	82	52	83.8	0.71
4	12_tanh_CANCRSRT_18in_all	80.5	54.3	82	0.71
5	20_10_2_sigma_CANCRSRT_18in_all	71.9	60.2	72.6	0.69

**Schemat 59** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka wystąpienia choroby nowotworowej uzyskane na danych pierwotnych (18 wejść), balansowanych ujemnie.

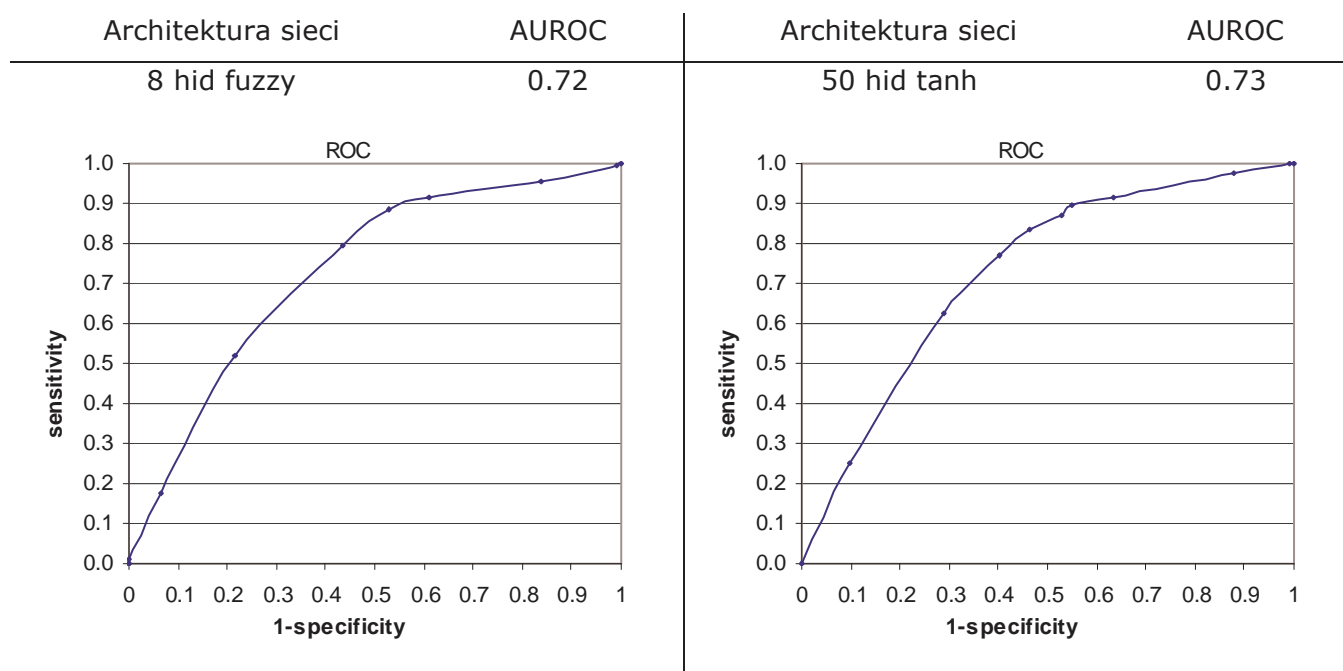




**Tabela 67** Najlepsze uzyskane wyniki modelowania ryzyka wystąpienia choroby nowotworowej z wykorzystaniem danych pierwotnych (18 wejść) – dane balansowane dodatnio.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1	8_fuzzy_CANCRSRT_18in_all	83.7	51.3	85.5	0.72
2	50_tanh_CANCRSRT_18in_all	83.1	51.3	84.9	0.73
3	10_7_tanh_CANCRSRT_18in_all	82.9	52	84.6	0.72
4	12_tanh_CANCRSRT_18in_all	81.5	52.9	83.1	0.72
5	18_8_sigma_CANCRSRT_18in_all	81	51.7	82.7	0.7

**Schemat 60** Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka wystąpienia choroby nowotworowej uzyskane na danych pierwotnych (18 wejść), balansowanych dodatnio.



## 7 Dyskusja wyników i zastosowania praktyczne

### 7.1 Zastosowanie sztucznych sieci neuronowych w farmakoekonomicie

Wykorzystywanie w farmakoekonomicie jako narzędzi modelowania efektu technik matematycznych, jest podejściem powszechnie akceptowanym. Zgodnie z polskimi wytycznymi modele powinny spełniać kilka warunków [103]. Wśród nich autorzy projektu wymieniają:

- prostotę (wyniki muszą być zrozumiałe dla badacza)
- przejrzystość prezentacji wyników
- konieczność przeprowadzenia analizy wrażliwości uzyskanych wyników pod kątem oceny mocy naukowej modelu
- otrzymane modele powinny być weryfikowane innymi technikami modelowania

Oprócz tego autorzy opracowania [135], wymieniają dodatkowy warunek modelu farmakoekonomicznego – możliwość podania jego postaci analitycznej.

Dodatkowym parametrem – oczywistym choć nie ujętym w wytycznych – jest skuteczność modelu, oceniana najlepszymi dostępnymi technikami analitycznymi.

W niniejszej pracy zastosowano sztuczne sieci neuronowe o różnej charakterystyce jako narzędzie modelowania wyniku terapeutycznego. Ich przydatność testowano wykorzystując informacje dotyczące terapii niedrobnokomórkowego raka płuc. Analiza możliwości ich zastosowania w farmakoekonomicie zostanie przeprowadzona zgodnie z powyższymi wytycznymi.

#### Prostota modelu oraz przejrzystość prezentacji wyników

Odpowiednio przygotowany i nauczony model neuronowy pracuje jako narzędzie podające na wyjściu wartość, będącą w ogólnej

charakterystyce wiernym odbiciem zmiennej modelowanej. W zależności od charakteru opracowywanego zagadnienia (np. medycznego) może to być zarówno odsetek pacjentów czy parametr epidemiologiczny jak i wartość kliniczna (wartość ciśnienia tętniczego krwi, stężenie glukozy czy długość okresu przeżycia). Dodatkowego podkreślenia wymaga fakt, że zasadnicza praca nauczonego modelu neuronowego odbywa się w sposób niewidoczny dla użytkownika.

W trakcie prac modelowano wartość definiowaną jako długość przeżycia w zależności od wdrożonej terapii, a więc wyjściem sieci jest konkretna zmienna obliczona dla interesującego badacza (ale także lekarza praktyka, pacjenta) wektora cech wejściowych. Prostota modelu jest więc zachowana w stopniu maksymalnym i niemożliwym do ulepszenia. Nie istnieje konieczność wykonywania żadnych dodatkowych operacji matematycznych, a wyniki są zawsze zrozumiałe na zainteresowanej nimi osoby. Pierwsze dwa wymogi są więc spełnione.

#### Analiza wrażliwości uzyskiwanych wyników pod kątem oceny mocy naukowej modelu

Powyższa analiza powinna być możliwie szeroka i obejmować zarówno różne możliwe kombinacje zmian wartości parametrów wejściowych (wieloczynnikowa) jak i różne zakresy zmian (wielokierunkowa). Jej celem jest weryfikacja odporności na dane wejściowe przypadkowe lub błędne, a także braku przypadkowości modelu. W przypadku modeli neuronowych schemat powyższej analizy obejmuje przygotowanie odpowiednio spreparowanych danych podawanych wybranemu systemowi na wejściu i analiza reakcji sieci (wartości wyjścia). W zależności od rozmiaru wektora wejściowego możliwa liczba kombinacji zmian wartości zmiennych wejściowych rośnie wykładniczo. Z tego powodu opisanej analizie poddano kilka najlepszych, wybranych modeli neuronowych.

W przypadku wyjścia o charakterze binarnym skuteczność testowanej terapii oceniano wykorzystując odsetek pacjentów przeżywających powyżej 35 tygodni. Wyjście ciągle umożliwia wykorzystanie średniego oraz maksymalnego okresu przeżycia jako miary skuteczności.

Kolejnym etapem jest przeprowadzenie analizy farmakoekonomicznej z wykorzystaniem zamodelowanych danych dotyczących efektywności, która pozwoli odpowiedzieć na pytanie dotyczące opłacalności stosowania poszczególnych terapii. Poniżej przedstawiono schemat tego typu analizy, z wykorzystaniem najlepszych uzyskanych modeli neuronowych, których charakterystykę przedstawiono w ramkach.

sieć 15_5hid sigma_30in_all/szum 5x5%/wyjście binarne/2 mln iteracji
--

- wpływ modyfikacji dawki chemioterapeutyków oraz leków wspomagających terapię na wynik leczenia

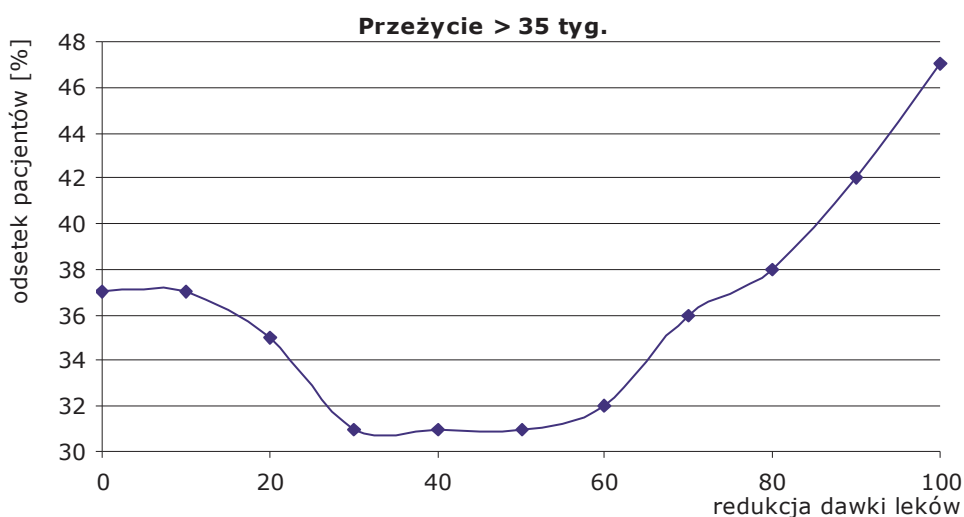
Doświadczenie ma na celu kontrolę zachowania się modelu neuronowego pod wpływem zmieniających się warunków zewnętrznych czyli przetestowanie jego wrażliwości. Schemat postępowania obejmował: przygotowanie zmodyfikowanych danych (w każdym z rekordów danych zmieniano krokowo, o 10% w każdym kroku dawki leków p/nowotworowych oraz wszystkich leków wspomagających – testowano *in-silico* reakcję wirtualnego pacjenta na zmienioną dawkę leku), pracę modelu oraz ocenę wyników definiowaną jako odsetek pacjentów przeżywających powyżej 35 tyg. Tabela 68 przedstawia reakcję systemu odzwierciedlającą okres przeżycia pacjentów na zmienioną (zredukowaną) dawkę chemioterapeutyku.

**Tabela 68** Zamodelowana reakcja systemu (okres przeżycia pacjentów) na zmienioną (zredukowaną) dawkę chemioterapeutyku.

Zmiana dawki leku [%]	Odsetek pacjentów przeżywających powyżej 35 tyg.	Odsetek pacjentów przeżywających poniżej 35 tyg.
0	37	63
10	37	63
20	35	65
30	31	69
40	31	69
50	31	69
60	32	68
70	36	64
80	38	62
90	42	58
100	47	53

Wykres 1 przedstawia przebieg zmian odsetka pacjentów w zależności od stopnia redukcji dawki leków.

**Wykres 1** Przebieg zmian odsetka pacjentów w zależności od stopnia redukcji dawki leków.



Powierzchniowa analiza powyższego wykresu może doprowadzić do wniosku, że system nie zachowuje się logicznie. O ile początkowy spadek skuteczności jest uzasadniony (zbyt mała dawka leków, w stosunku do

potrzeb chorego organizmu), o tyle druga faza wykresu budzi zastrzeżenia. Jeśli jednak powyższy wykres zestawiony zostanie z wynikami doświadczeń badających wpływ rodzaju opieki nad pacjentem w terminalnym stadium choroby nowotworowej, a długością życia - interpretacja wyników systemu może okazać się zupełnie inna. Jednym z zarzucanych - choć nadal pozostających w zakresie tematów kontrowersyjnych - schematów opieki jest zapewnienie najlepszej możliwej opieki, jednak bez stosowania środków chemioterapeutycznych (*ang. BSC - Best Supportive Care*). Analiza doniesień naukowych z tej dziedziny dowodzi, że metoda BSC - choć zarzucana - stanowi pewną alternatywę dla agresywnej, paliatywnej chemioterapii, a rezultaty medyczne dowodzą, że może być ona uznawana za porównywalną. W świetle powyższych rozważań, reakcja sieci na redukcję dawki leków staje się wytłumaczalna i stanowić może kolejny przyczynek do dyskusji nad racjonalnością stosowania metody BSC.

- wpływ zamiany chemioterapeutyków na wyniki terapii

Jedną z kolejnych wykonanych prób kontroli wrażliwości modelu neuronowego, jest sprawdzenie jego reakcji w odniesieniu do danych literaturowych dotyczących skuteczności leków p/nowotworowych. W tym celu przygotowano spreparowane dane, w których zastępowano sumaryczną dawkę jednego leku - innym. Jest to sytuacja symulująca codzienną praktykę kliniczną. Na podstawie stanu klinicznego pacjenta, najlepszej wiedzy oraz w granicach dostępnych środków finansowych, na drodze pewnego algorytmizowanego procesu decyzyjnego lekarz prowadzący terapię, dokonuje wyboru schematu leczenia. Przykładem sytuacji, gdzie wśród możliwych alternatyw należy zdecydować się na jedną z nich jest sytuacja, w której dla pacjenta w nieoperacyjnym stadium niedrobnokomórkowego raka płuc, są do wyboru dwie możliwości: schemat Gemcytabina - Cis-platyna oraz schemat Winorelbina - Cis-platyna.

Przygotowanie danych dla potrzeb doświadczenia numerycznego polegało na zastąpieniu gemcytabiny odpowiadającą jej dawką

winorelbiny. Tabela 69 przedstawia porównanie wyników z danych pierwotnych (z gemcytabiną, bez winorelbiny) oraz zamodelowanych (gemcytabina zastąpiona winorelbina).

**Tabela 69** Porównanie skuteczności chemioterapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (gemcytabina) oraz zamodelowanych (winorelbina).

	Liczba pacjentów przeżywających	
	>35 tyg.	<35 tyg.
Dane pierwotne	6	20
Dane zamodelowane	6	20

Powyższa tabela zawiera dane potwierdzające informacje z niezależnych doniesień naukowych, które pokazują, że różnica w skuteczności między gemcytabiną oraz winorelbina w tym typie nowotworu jest nieistotna statystycznie.

Przeprowadzono jednocześnie doświadczenie odwrotne, polegające na zastąpieniu dawki winorelbiny odpowiadającą jej dawką gemcytabiny. Postępowanie takie miało na celu zbadanie wrażliwości modelu. Podobnie jak powyżej również w tym przypadku model nie stwierdził istotnych różnic w skuteczności obu terapii (Tabela 70).

**Tabela 70** Porównanie skuteczności chemioterapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (winorelbina) oraz zamodelowanych (gemcytabina).

	Liczba pacjentów przeżywających	
	>35 tyg.	<35 tyg.
Dane pierwotne	16	24
Dane zamodelowane	16	24

Powyższa sytuacja sugeruje typ analizy farmakoekonomicznej możliwej do zastosowania. Zakładając identyczność wyników można wykorzystać analizę minimalizacji kosztów. Dane dotyczące kosztów stosowania poszczególnych schematów chemioterapii zaczerpnięto z pracy [120]. Pod uwagę wzięto koszty całkowite na jednego pacjenta,

uwzględniając wszystkie elementy składowe. Porównano zarówno koszty leczenia ambulatoryjnego jak i szpitalnego, wykorzystując opcje najkorzystniejsze pod kątem finansowym dla poszczególnych metod.

**Tabela 71** Porównanie kosztów całkowitych dla schematów cis-platyna/winorelbina oraz cis-platyna/gemcytabina [120].

		Schemat		
		cis-platyna winorelbina	cis-platyna gemcytabina	
Koszt	średni/pacjenta	ambulatoryjne	3043 ( $\pm$ 2384)	12472 ( $\pm$ 7 287)
		szpitalne	9678 ( $\pm$ 4 383)	16919 ( $\pm$ 8 778)
	średni/cykl	ambulatoryjne	864 ( $\pm$ 471)	80842 ( $\pm$ 3 305)
		szpitalne	3755 ( $\pm$ 2 310)	3224 ( $\pm$ 1 794)

Tabela 72 zawiera średnie ceny porównywanych cytostatyków we wszystkich ośrodkach, z których pochodziły dane wykorzystane w trakcie badania.

**Tabela 72** Średnie ceny porównywanych cytostatyków we wszystkich ośrodkach.

	Średnia cena leku [zł]
Navelbina amp. 50 mg x 1	544.40
Navelbina amp. 10 mg x 1	119.53
Gemcytabina amp. 1000 mg x 1	782.27
Gemcytabina amp. 200 mg x 1	164.00

Analiza danych, które zawiera Tabela 71 oraz Tabela 72, wskazuje na efektywność kosztową schematu z winorelbina. Po jej zastąpieniu droższą gemcytabiną - przy braku różnic w efektach - koszty całkowite będą wyższe.

- wpływ zmiany chemioterapeutyków na skuteczność terapii

Kontynuując powyższe rozważania, przeprowadzono analizę *in-silico* wpływu zamiany drugiego chemioterapeutyku podawanego w schemacie z lekiem podstawowym. Wśród dostępnych danych opisujących pacjentów znajdowały się dwa leki – pochodne platyny, łączone z etopozydem, gemcytabiną i winorelbina. Cis-platyna oraz karboplatyna nie wykazują



istotnej różnicy w skuteczności. Jednocześnie jednak karboplatyna jest lekiem obciążonym mniej nasilonymi działaniami niepożądanymi (nefrotoksyczność), natomiast różnica cen na jej niekorzyść powoduje, że jest wykorzystywana rzadziej niż cis-platyna.

Powstaje więc pytanie, na które odpowiedzieć może jedynie przeprowadzenie analizy farmakoekonomicznej. Czy mimo wyższej ceny opłacalnym byłoby stosowanie karboplatyny ze względu na niższe koszty terapii działań toksycznych.

Tabela 73 przedstawia wyniki modelowania zamiany dawki karboplatyny na odpowiadającą jej dawkę cis-platyny.

**Tabela 73** Porównanie skuteczności chemioterapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (karboplatyna) oraz zamodelowanych (cis-platyna).

	Liczba pacjentów przeżywających	
	>35 tyg.	<35 tyg.
Dane pierwotne	4	6
Dane zamodelowane	6	4

Jednocześnie wykonano doświadczenie odwrotne – zamieniono dawkę cis-platyny na odpowiadającą jej dawkę karboplatyny.

**Tabela 74** Porównanie skuteczności chemioterapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (cis-platyna) oraz zamodelowanych (karboplatyna).

	Liczba pacjentów przeżywających	
	>35 tyg.	<35 tyg.
Dane pierwotne	33	57
Dane zamodelowane	23	67

W pierwszym przypadku system wskazuje na możliwość niższej skuteczności karboplatyny w porównaniu z zamodelowaną skutecznością leku odnośnikowego (cis-platyny). Drugi etap doświadczenia potwierdza powyższą hipotezę.

Analiza danych przedstawionych w powyższych tabelach pokazuje elastyczność i konsekwencję działania wyselekcjonowanego systemu neuronowego. Jednocześnie pokazuje możliwość wykorzystania modeli neuronowych jako narzędzi wstępnej analizy pod kątem zastosowania w analizie farmakoekonomicznej.

- wpływ zmiany leków wspomagających terapię (na przykładzie antiemetyków) na skuteczność terapii

Etopozyd podobnie jak i gemcytabina oraz winorelbina wykazują stosunkowo niewielki potencjał emetyczny. Jednocześnie jednak nie stosuje się ich w monoterapii, a leki z którymi są najczęściej łączone w schematach – cis-platyna oraz carboplatyna – wykazują bardzo silne działanie emetyczne. Ryzyko wystąpienia wymiotów jest tak duże, że najczęściej leki p/wymiotne podawane są profilaktycznie przed rozpoczęciem wlewu leku z grupy pochodnych platyny. Najczęściej stosowanymi preparatami są leki z grupy antagonistów receptora 5HT<sub>3</sub> – ondansetron, tropisetron czy granisetron (niezarejestrowany w Polsce). Są to preparaty skuteczne jednak bardzo drogie, obciążające konto płatnika za usługi medyczne.

Lekiem z innej grupy również stosowanym w podobnych sytuacjach jest metoklopramid – ośrodkowy i obwodowy antagonist receptorów dopaminowych, zwiększający również wrażliwość receptorów cholinergicznym. Jest on jednocześnie kilkadziesiąt razy tańszy niż antagoniści receptora dla serotoniny. Pytaniem, które może się w takiej sytuacji nasunąć jest różnica w skuteczności obu grup i ich wpływ na wyniki leczenia.

Aby odpowiedzieć na to pytanie przeprowadzono testy *in-silico* polegające na zamianie jednego leku drugim i ocenie efektu podobnej operacji. W pierwszym kroku dawkę ondansetronu zastąpiono odpowiadającą jej dawką metoklopramidu. Jak wynika z danych, które przedstawia Tabela 75 reakcja systemu na zmianę leku sugeruje wzrost

skuteczności wyrażonej w postaci odsetka pacjentów przeżywających powyżej 35 tygodni.

**Tabela 75** Porównanie wpływu zmiany leku antiemetycznego na wynik terapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (ondansetron) oraz zamodelowanych (metoklopramid).

	Liczba pacjentów przeżywających	
	>35 tyg.	<35 tyg.
Dane pierwotne	28	50
Dane zamodelowane	43	35

W celu sprawdzenia wrażliwości modelu przeprowadzono doświadczenie analogiczne w stosunku do przedstawionego wcześniej. Wykonano operację odwrotną tzn. dawkę metoklopramidu zastąpiono odpowiadającą jej dawką ondansetronu. Tabela 76 obrazuje reakcję systemu na powyższą operację.

**Tabela 76** Porównanie wpływu zmiany leku antiemetycznego na wynik terapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (metoklopramid) oraz zamodelowanych (ondansetron).

	Liczba pacjentów przeżywających	
	>35 tyg.	<35 tyg.
Dane pierwotne	13	12
Dane zamodelowane	7	18

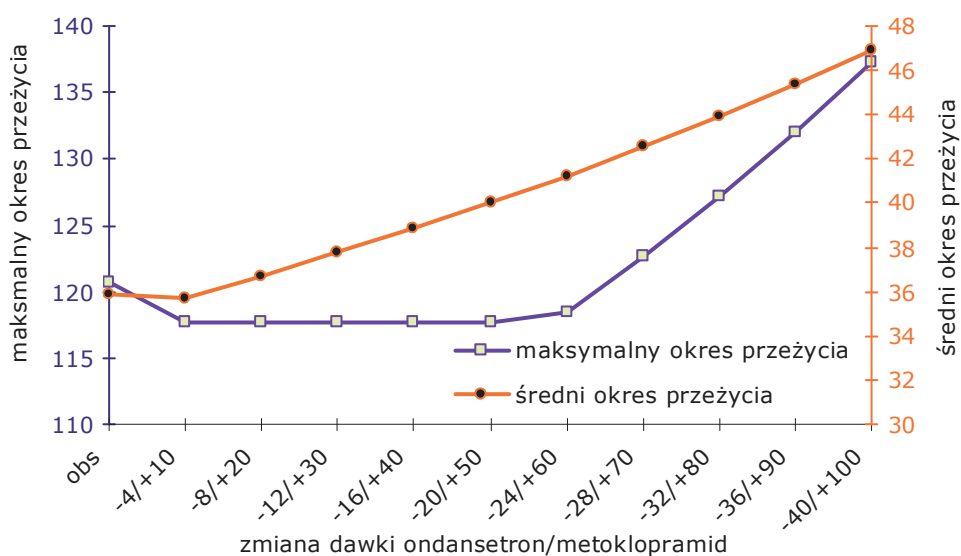
Analiza danych przedstawionych w powyższych tabelach dowodzi elastyczności i konsekwencji działania wyselekcjonowanego systemu neuronowego. Jednocześnie może być przyczynkiem do dyskusji i początkiem analizy farmakoekonomicznej mającej na celu oszacowanie opłacalności stosowania w tym konkretnym przypadku drogich preparatów z grupy blokerów receptora serotoninowego.

- wpływ zmiany leków wspomagających terapię (na przykładzie antiemetyków) na skuteczność terapii

Doświadczenie polegało na stopniowej zmianie dawki leku p/wymiotnego z grupy blokerów receptora serotoninowego (5HT<sub>3</sub>) odpowiadającą jej dawką metoklopramidu.

Testowano wpływ powyższej operacji na średni okres przeżycia pacjentów.

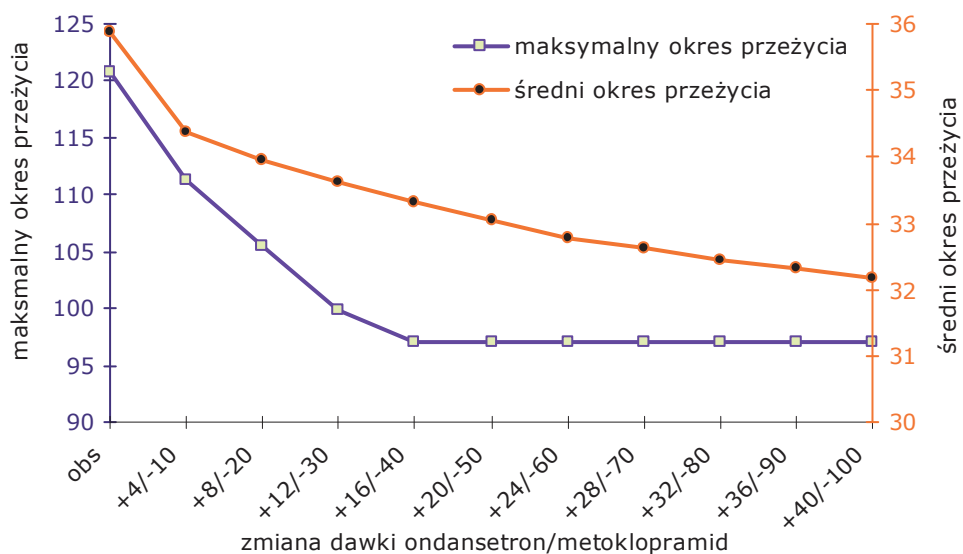
**Wykres 2** Zamodelowany przebieg zmian okresu przeżycia pacjentów w zależności od charakteru środka p/wymiotnego.



Testowany system neuronowy przewiduje, że zmiana leku antiemetycznego nie będzie miała negatywnych konsekwencji w odniesieniu do efektu farmakoterapii. Charakter zmian zaproponowanych przez sieć neuronową, przemawia na korzyść testowanego metoklopramidu.

Analogicznie wykonano drugą część doświadczenia, zamieniając dawkę metoklopramidu odpowiadającą jej dawką ondansetronu oraz testując *in-silico* reakcję pacjentów na podobną zmianę.

**Wykres 3** Zamodelowany przebieg zmian okresu przeżycia pacjentów w zależności od charakteru środka p/wymiotnego.



Analizując powyższe wykresy zauważyć można logikę i konsekwencję działania sieci neuronowej, która objawia się w odmiennej reakcji na podawane przeciwstawne dane.

### Analiza uzyskanych wyników oraz weryfikacja modeli neuronowych innymi technikami modelowania

Dobór odpowiednich metod odniesienia zależy od charakteru opracowywanego problemu. W trakcie prac przyjęto następującą metodykę: w przypadku modelowania zmiennej o charakterze ciągłym budowano porównawcze modele o charakterze liniowym, natomiast w przypadku zmiennej binarnej najbardziej odpowiednią techniką jest regresja logistyczna. Najlepsze modele neuronowe dla poszczególnych etapów pracy porównywano z odpowiednimi modelami regresyjnymi stosując odpowiednie dla charakteru zmiennej wyjściowej parametry.

a) Modelowanie z wykorzystaniem danych pierwotnych – wyjście ciągłe

Najlepszy uzyskany na tym etapie model neuronowy dał wynik 20.97 błędu RMSE (Rozdział 5.2.1). Nie jest to wartość zadowalająca, która pozwalałaby na zastosowanie modelu w praktyce. Przyczyną takiego stan

rzeczy w opinii autora jest mała ilość rekordów zbioru pierwotnego i – co jest tego bezpośrednią konsekwencją – zbiorów uczących wykorzystywanych w trakcie 10-cio krotnego wzajemnego sprawdzania. Zbiór zawierający 100 rekordów danych, przy tak skomplikowanym i wielowymiarowym problemie, nie może być wystarczający. Dodatkowo ciągły charakter zmiennej zależnej powoduje, że problem stanowi zbyt duże wyzwanie dla systemu. Analiza odpowiedzi nauczonej sieci (dla najlepszego modelu 2hid M-SLIDE – Rozdział 5.2.1) pokazuje, że system nie potrafi odwzorować wartości skrajnych, w tym przypadku przekraczających 100 tygodni przeżycia.

Model oparty o regresję liniową osiąga wartość błędu RMSE 22.79, a więc wyższy niż model neuronowy, choć nie jest to różnica istotna statystycznie.

b) Modelowanie z wykorzystaniem danych przeformatowanych

– Dane zaszumione - wyjście ciągłe

Kolejny krok to testowanie wpływu zaszumiania danych (różne parametry szumu) na otrzymywane wyniki, a bezpośrednim celem było znalezienie modelu zdolnego zejść poniżej wartości 20 błędu RMSE dla testu generalizacji (Rozdział 5.2.2.1). Na tym etapie nie zmieniano parametrów zmiennej zależnej. W trakcie poszukiwań architektury i parametrów sieci udało się wyselekcjonować model, dla którego wartość błędu RMSE zmalała do wartości 20.6, wykorzystując zbiór zaszumiany z parametrami 5x5%. Sieć miała identyczną architekturę jak poprzednia (2hid M-SLIDE). Ze względu na to, że różnica nie jest istotna statystycznie, a koszt obliczeniowy większy (ze względu na większą ilość rekordów zaszumionego zbioru uczącego) nie jest to alternatywa dominująca. Błąd modelu liniowego zbudowanego z wykorzystaniem danych zaszumianych w wyniku testu zdolności generalizacji osiągnął wartość 22.76.

- Zmienna wyjściowa dyskretna

Na tym etapie miarą dopasowania modelu był bezwzględny odsetek przypadków prawidłowo zaklasyfikowanych do poszczególnych klas. W przypadku obydwu testowanych wartości testowano dwie wartości szerokości przedziału – 15 oraz 35 tygodni.

Zarówno w przypadku podziału na trzy jak i cztery klasy, lepsze wyniki całkowite osiągnięto przy 35-cio tygodniowej długości przedziału - 67% w obydwu przypadkach (Rozdział 5.2.2.2). Jednocześnie jednak cechą charakterystyczną jest nierówny rozkład wyników w zależności od klasy. Zjawisko to nie występuje w przypadku zastosowania 15-to tygodniowej szerokości przedziału, jednak wyniki testowania generalizacji nie są zadowalające - najlepsze modele wykazały odpowiednio jedynie 54% i 38% skuteczność. Najprawdopodobniej przyczyną przedstawionej powyższej sytuacji, jest nierównomierny rozkład rekordów w poszczególnych klasach (przedział 35-cio tygodniowy) oraz zbyt mała ilość rekordów zbioru bazowego w obu testowanych przypadkach. Uniemożliwia to uzyskanie wystarczająco skutecznego modelu nawet wykorzystując tak przekształcony zestaw atrybutów. Korzystając z zestawu danych zawierających znacząco więcej rekordów w jednej z klas, sieć – zachowując się bardzo logicznie – uczy się rozpoznawać głównie tą klasę, „przewidując”, że będzie ona w znaczący sposób rzutować na wyniki całkowite. Z drugiej strony przygotowanie zbalansowanego zestawu klas mimo poprawy wyników poszczególnych z nich, skutkuje pogorszeniem rezultatów całkowitych (wspomniana powyżej mała ilość rekordów).

Z punktu widzenia praktyki zastosowanie modelu ze zbalansowanymi klasami nie jest możliwe. Wykorzystanie modeli z 35-cio tygodniowymi przedziałami skutkowałoby ostrożnością modelu i klasyfikowaniem rekordów do klasy pacjentów przeżywających najkrótszy okres.

Wyniki uzyskane w trakcie wcześniejszych etapów nie pozwalają na zastosowanie proponowanego systemu w praktyce codziennej. W związku z tym zaprojektowano kolejny krok, mający na celu przetestowanie zdolności generalizacji sieci w opisywanym problemie, przy wykorzystaniu

danych o binarnym charakterze wyjścia. W przypadku binarnej zmiennej zależnej standardowym porównawczym narzędziem modelowania jest regresja logistyczna. Dla każdej z testowanych opcji preprocessingu danych przygotowano porównawcze modele logistyczne oraz dodatkowo liniowe, które zestawiano z najlepszymi modelami neuronowymi.

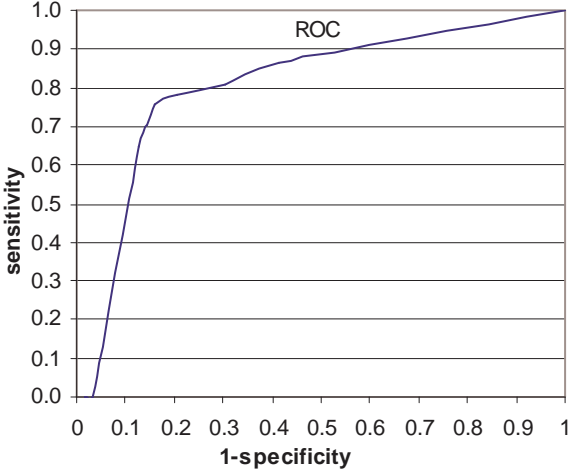
- dane pierwotne

Tabela 77 oraz Tabela 78 przedstawiają porównanie parametrów oceny modeli dla najlepszego systemu neuronowego oraz statystycznego (przy wartości progu 35 tygodni).

**Tabela 77** Porównanie najlepszego modelu neuronowego z modelami regresyjnymi - dane pierwotne z wyjściem binarnym (próg 35 tygodni)

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	20_10_2hid fsr_30in_all	81	76	84	0.81
2.	Log_regression	80	86	75	0.87
3.	Lin_regression	77	62	86	0.81

**Tabela 78** Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych pierwotnych z wyjściem binarnym (próg 35 tygodni) – różne metody modelowania.

Architektura najlepszej sieci	AUROC	Wykres ROC
20_10_2hid fsr_30in_all	0.81	



Log_regression	0.87	
Lin_regression	0.81	

Model neuronowy osiągnął najlepszy wynik w teście generalizacji, choć różnice nie są istotne statystycznie. Istotnym faktem jest, że model logistyczny uzyskano wykorzystując jedynie trzy pary zbiorów testowych, ponieważ przy przyjętych założeniach dla pozostałych par metoda nie była zbieżna (nie udało się uzyskać parametrów funkcji). Nie można więc bezpośrednio porównywać modelu neuronowego (dziesięć par – 10-cio krotne wzajemne sprawdzanie) oraz logistycznego (trzykrotne wzajemne sprawdzanie).

- dane zaszumione

Tabela 79 oraz Tabela 80 przedstawiają porównanie wyników trzech dróg modelowania dla danych zaszumianych z parametrami 2x2%. Podobnie jak wcześniej, również w tym przypadku nie udało się

wykorzystać wszystkich par testowy-uczący w przypadku regresji logistycznej (8-mio krotne wzajemne sprawdzanie).

**Tabela 79** Porównanie najlepszego modelu neuronowego z modelami regresyjnymi - dane zaszumiane (parametry szumu - 2x2%) z wyjściem binarnym (próg 35 tygodni).

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	20_16_8_4hid sigma_30in_all	82	73	87	0.78
2.	Log_regression	74	73	74	0.73
3.	Lin_regression	80	68	87	0.75

**Tabela 80** Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych zaszumianych (2x2%) z wyjściem binarnym (próg 35 tygodni) - różne metody modelowania.

Architektura najlepszej sieci	AUROC	Wykres ROC
20_16_8_4hid sigma_30in_all	0.78	

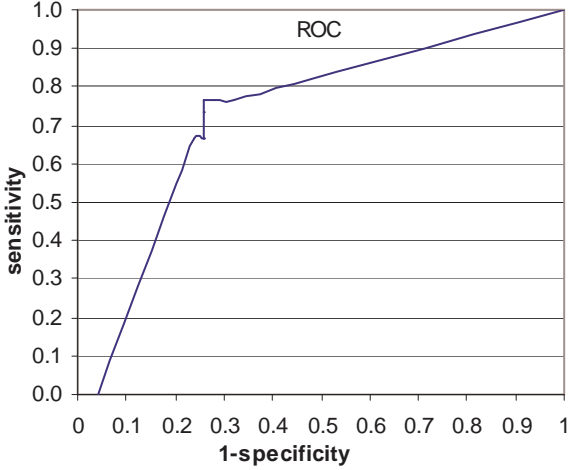
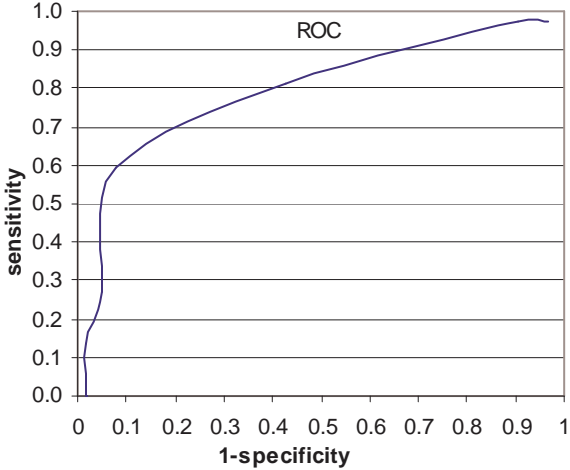
Log_regression	0.73	
Lin_regression	0.75	

Tabela 81 oraz Tabela 82 zawierają analogiczne wyniki dla danych zaszumianych z parametrami 5x5%. Dodatkowo dla najlepszej uzyskanej sieci przeprowadzono doświadczenie polegające na zmianie sposobu prezentacji danych. Zastosowana technika *leave-one-out* jest specyficznym typem metody k-krotnego wzajemnego sprawdzania. Zastosowano system wyłączenia pojedynczego rekordu do jednoskładnikowego zbioru testowego. Pozostałe rekordy stanowiły zbiór uczący.

**Tabela 81** Porównanie najlepszego modelu neuronowego z modelami regresyjnymi - dane zaszumiane (parametry szumu – 5x5%) z wyjściem binarnym (próg 35 tygodni).

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	15_5hid sigma_30in_all	84	76	89	0.82
2.	15_5hid sigma_leave-one-out	72	57	81	0.73
3.	Log_regression	73	70	74	0.7
4.	Lin_regression	77	65	84	0.81

**Tabela 82** Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych zaszumianych (5x5%) z wyjściem binarnym (próg 35 tygodni) – różne metody modelowania.

Architektura najlepszej sieci	AUROC	Wykres ROC
15_5hid sigma_30in_all	0.82	

15_5hid sigma_leave-one-out	0.73	
Log_regression	0.7	
Lin_regression	0.81	

Podobnie jak w przypadku danych pierwotnych, modele neuronowe uzyskane na danych zaszumionych również okazały się być skuteczniejsze niż systemy odnośnikowe.

Dodatkowa weryfikacja modelu z wykorzystaniem techniki *leave-one-out* pozwala na potwierdzenie właściwości predykcyjnych modelu.

### Postać analityczna modelu

Podanie postaci analitycznej modelu farmakoekonomicznego, jest jednym z warunków wymienianych w odpowiednich wytycznych. Pośrednim celem niniejszej pracy jest falsyfikacja tego założenia. Cel ten w opinii autora osiągnąć można dowodząc, że modele w postaci niejawnej (np. modele neuronowe), są co najmniej równie przydatne jak modele jawne. Przedstawione powyżej wyniki dowodzą, że modele neuronowe mogą być z powodzeniem stosowane jako narzędzia modelowania w farmakoekonomice, mimo charakteru „czarnej skrzynki”.

Jedną z możliwości uniknięcia podobnych zarzutów jest zastosowanie sieci rozmytych, dla których istnieją półempiryczne metody analizy.

## 7.2 Zastosowanie sztucznych sieci neuronowych w epidemiologii

Jedną z głównych cech epidemiologicznych zbiorów danych jest wielowymiarowość. Problem ten stanowi jedno z głównych ograniczeń przy stosowaniu do ich analizy klasycznych modeli statystycznych. Jest to jednocześnie jeden z powodów, dla których wykorzystano sztuczne sieci neuronowe potrafiące ominąć tę niedogodność. Cel pracy czyli dowiedzenie, że sieci neuronowe mogą skutecznie wykorzystywać tego typu zbiory został osiągnięty, o czym świadczą uzyskiwane na poszczególnych etapach wyniki.

Kolejną cechą charakterystyczną dla zbiorów informacji o charakterze epidemiologicznym, jest możliwość występowania rekordów fałszywych, sprzecznych lub zaszumionych. W przypadku technik statystycznych stosowane są metody ich eliminacji (np. usuwanie danych znacznie odbiegających od wartości średniej). Zastosowanie odpowiednio dobranej architektury i parametrów uczenia sieci neuronowej umożliwia -

podobnie jak powyżej – uzyskanie modelu zdolnego pracować z tego typu danymi, bez stosowania dodatkowych technik.

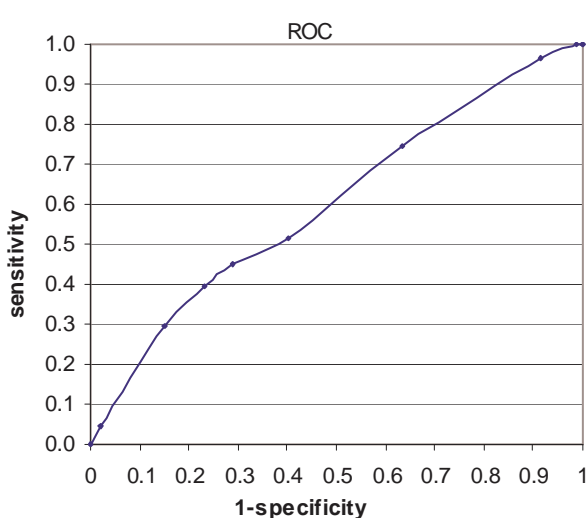
### Modelowanie ryzyka wystąpienia cukrzycy ciężarnych

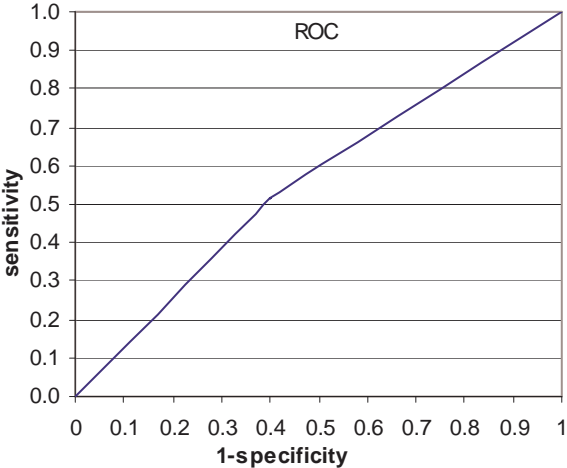
Najlepsze wyniki modelowania neuronowego (sieć 20\_18\_10hid fsr, pracującą w oparciu o ośmioskładnikowy wektor wejściowy z danymi balansowanymi dodatnio), uzyskane w tej części prac (Rozdział 6.2.3) zostały porównane z techniką odnośnikową – modelowaniem z wykorzystaniem techniki regresji logistycznej.

**Tabela 83** Porównanie - najlepsze uzyskane wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych przekodowanych – różne metody modelowania.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	20_18_10hid fsr_8in_dec_bal+	71.1	72.1	45.1	0.60
2.	Log_regression	59.8	51.6	60.1	0.56

**Schemat 61** Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych przekodowanych – różne metody modelowania (cukrzyca ciężarnych, 8 wejść).

Architektura najlepszej sieci	AUROC	Wykres ROC
20_18_10hid fsr_dec_bal+	0.60	

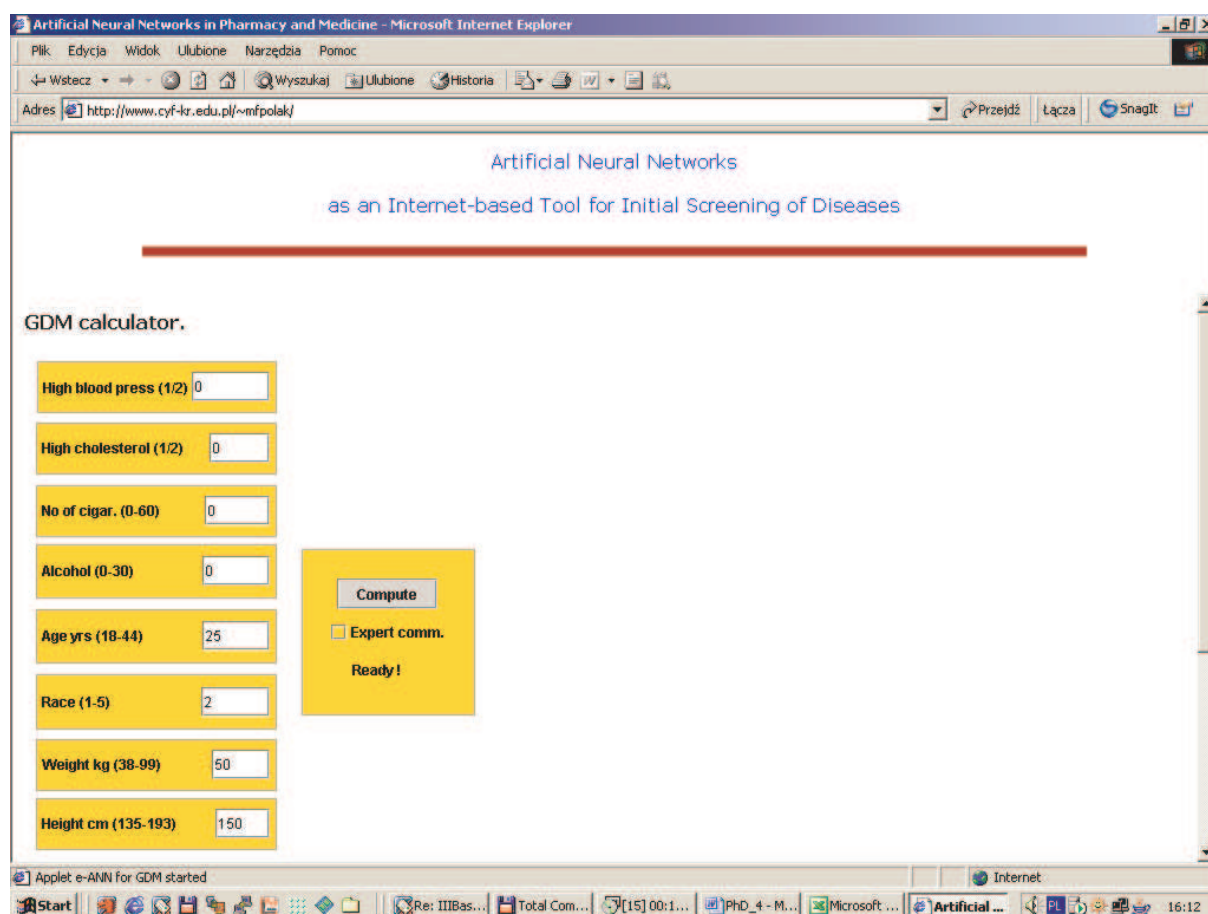
Log_regression	0.56	
----------------	------	--

Analiza wyników zawartych w powyższej tabeli dowodzi, że korzystając z modelu neuronowego można osiągnąć lepsze rezultaty niż w oparciu o techniki statystyczne. Powyższy model nie może służyć do diagnostyki cukrzycy ciężarnych. Błąd, którym jest obarczone jego działanie wciąż nie umożliwia podobnych doświadczeń. Niemniej jednak jego planowane przeznaczenie jest inne – ma stanowić narzędzie wstępnego skriningu kobiet ciężarnych, pod kątem konieczności wykonywania pełnych badań diagnostycznych. Stąd czułość systemu ma znaczenie nadrzędne w stosunku do jego specyficzności.

Wcześniejsza wersja modelu neuronowego uzyskana w oparciu o 12 zmiennych sterujących posłużyła do zbudowania wzmiankowanego powyżej systemu skriningu opartego o powszechnie dostępne narzędzie internetowe napisane w języku Java, dla którego silnikiem był najlepszy uzyskany na ówczesnym etapie system neuronowy [101]. Rysunek 1 przedstawia zrzut ekranowy części systemu, który dostępny jest pod adresem internetowym <http://www.cyf-kr.edu.pl/~mfpolak>.



**Rysunek 1** Zrzut ekranowy internetowego systemu predykcji cukrzycy ciężarnych pracującego w oparciu o model neuronowy (<http://www.cyf-kr.edu.pl/~mfpolak/>).



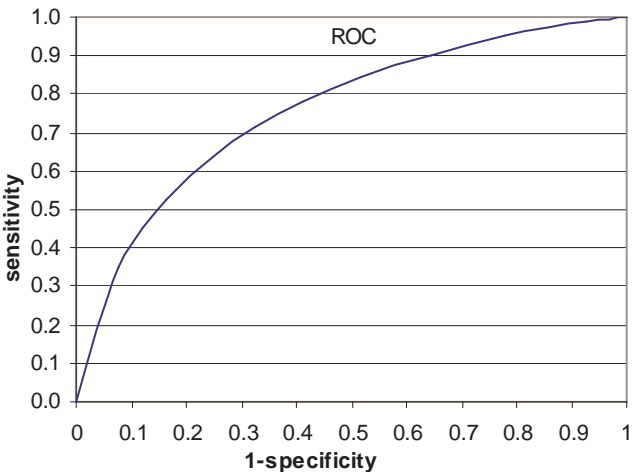
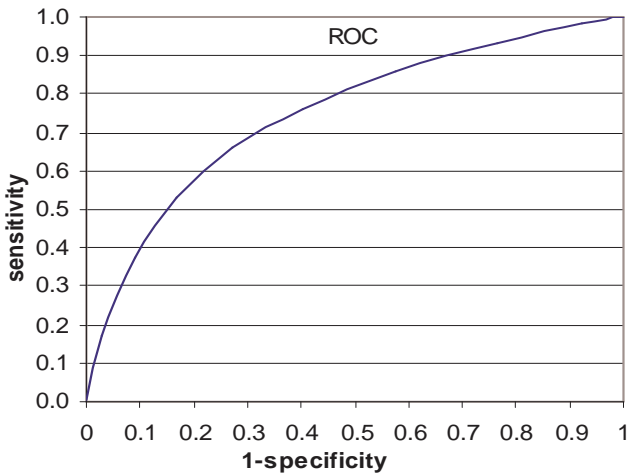
## Modelowanie ryzyka wystąpienia nadciśnienia tętniczego

Zgodnie z przyjętym schematem najlepsze wyniki otrzymane na drodze modelowania neuronowego (sieć 2hid tanh, pracującą w oparciu o sześćskładnikowy wektor wejściowy z danymi balansowanymi dodatnio – Rozdział 6.3.3) porównano z wynikami uzyskanymi technikami statystycznymi.

**Tabela 84** Porównanie - najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego – różne metody modelowania.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	2hid tanh_6in_all_bal+	71.27	62.99	74.66	0.76
2.	Log_regression	70.94	64.63	73.53	0.76

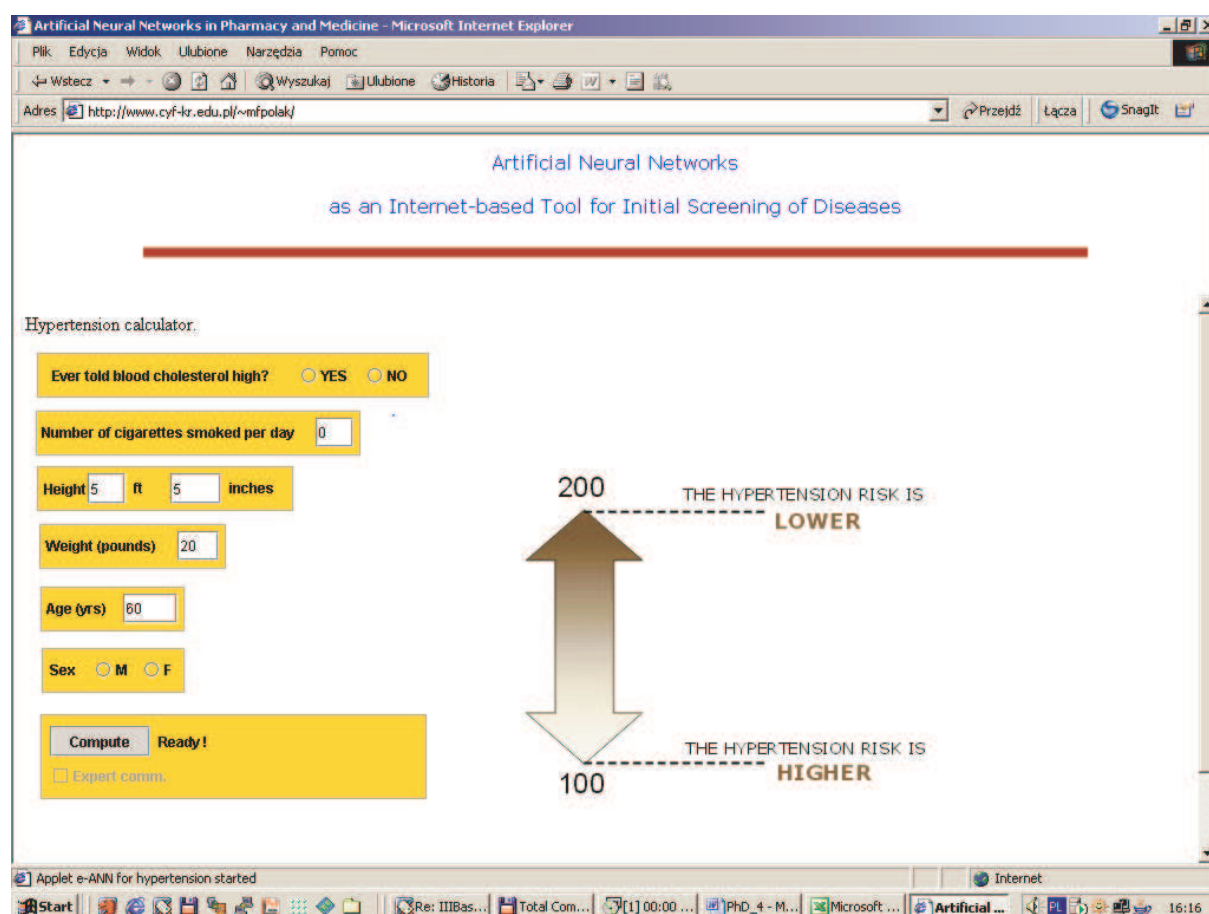
**Schemat 62** Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych przekodowanych – różne metody modelowania (nadciśnienie tętnicze, 6 wejść).

Architektura najlepszej sieci	AUROC	Wykres ROC
2hid tanh_6in_all_bal+	0.76	
Log_regression	0.76	

Najlepszy model neuronowy uzyskany w oparciu o 6 zmiennych sterujących posłużył do zbudowania systemu skriningu opartego o powszechnie dostępne narzędzie internetowe napisane w języku Java [102]. Rysunek 2 przedstawia zrzut ekranowy części systemu, który dostępny jest pod adresem internetowym <http://www.cyf-kr.edu.pl/~mfpolak>.

Podobnie jak narzędzie prezentowane powyżej system ten nie ma na celu zastąpienie klasycznej procedury diagnostycznej. Może być jednak wykorzystany jako system skriningowy oraz fragment działań w zakresie promocji zdrowia. Zwracając uwagę pacjentów na problem nadciśnienia i wskazując czynniki ryzyka może przyczynić się do spopularyzowania wiedzy na temat tego schorzenia.

**Rysunek 2** Zrzut ekranowy internetowego systemu predykcji nadciśnienia tętniczego pracującego w oparciu o model neuronowy (<http://www.cyf-kr.edu.pl/~mfpolak/>).



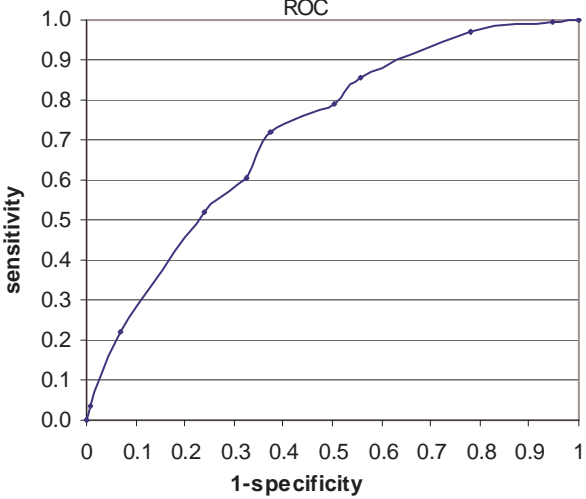
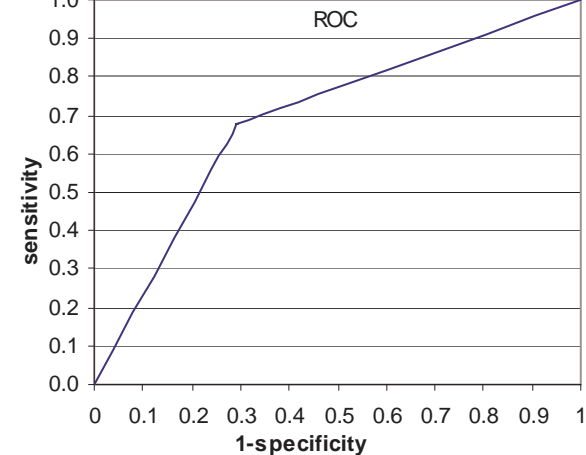
### Modelowanie ryzyka ostrego incydentu wieńcowego

Najlepszy model neuronowy (sieć 3hid sigma, pracującą w oparciu o wektor wejściowy z 28 składnikami, z danymi balansowanymi dodatnio – Rozdział 6.4.3) porównano z wynikami regresji logistycznej.

**Tabela 85** Porównanie - najlepsze uzyskane wyniki modelowania ryzyka ostrego incydentu wieńcowego – różne metody modelowania.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	3 sigma_HEARTATK_28in_bal+	70.14	62.59	72.01	0.71
2.	Log_regression	68.4	70.7	67.8	0.69

**Schemat 63** Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych pierwotnych – różne metody modelowania (ostry incydent wieńcowy, 28 wejść).

Architektura najlepszej sieci	AUROC	Wykres ROC
3sigma_HEARTATK_28in_bal+		
Log_regression	0.69	

### Modelowanie ryzyka wystąpienia choroby nowotworowej

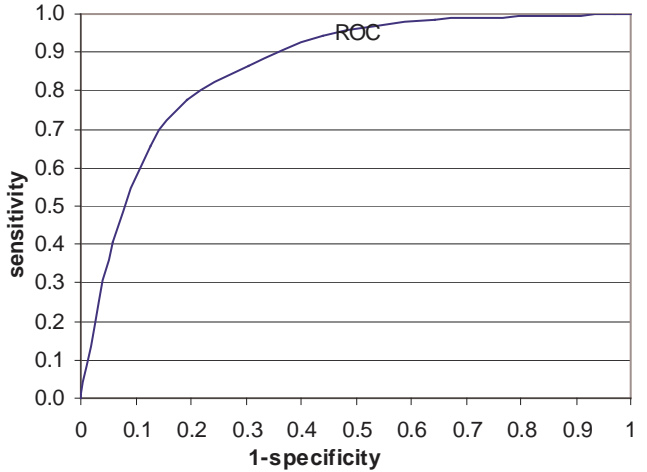
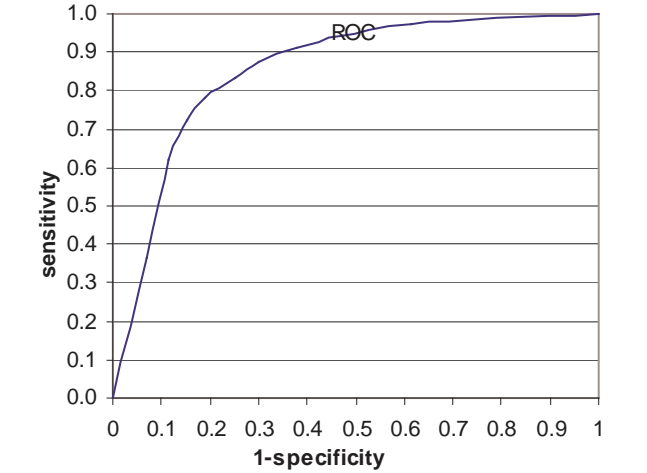
Dla opisywanego zagadnienia lepsze rezultaty osiągnięto korzystając z danych opisanych przez 101 zmiennych sterujących (Rozdział 6.5.3). Większy rozmiar wektora wejściowego wiązał się z możliwością zbudowania modeli neuronowych charakteryzujących się wyższą zdolnością generalizacji. Redukcja jego rozmiaru do 18 wejść skutkowałą

obniżeniem skuteczności. Praktyczne zastosowanie systemu opartego o tak dużą liczbę koniecznych informacji jest ograniczone. Biorąc pod uwagę, że zbadana różnica w skuteczności obu systemów jest istotna statystycznie w dalszych rozważaniach wykorzystywano najlepszą sieć opartą o bardziej rozbudowany wektor wejściowy.

**Tabela 86** Porównanie - najlepsze uzyskane wyniki modelowania ryzyka wystąpienia choroby nowotworowej – różne metody modelowania.

Nr	Architektura sieci	GLOBAL	Odsetek prawidłowych odpowiedzi dla klas		AUROC
			1	0	
1.	20hid MSLIDE_CANCRSRT_101in_bin	82.6	74.5	83.1	0.86
2.	Log_regression	82.9	74.7	83.3	0.85

**Schemat 64** Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych pierwotnych – różne metody modelowania (choroba nowotworowa, 101 wejść).

Architektura najlepszej sieci	AUROC	Wykres ROC
<p style="text-align: center;">20hid MSLIDE_CANCER_101in</p>	0.86	
<p style="text-align: center;">Log_regression</p>	0.85	

## 8 Wnioski

1. Sztuczne sieci neuronowe mogą być cennym narzędziem modelowania w farmakoekonomice, czego dowodzą wyniki osiągnięte nawet na ograniczonym zbiorze danych.
2. Mimo braku możliwości podania analitycznej postaci modelu dla części modeli neuronowych – należy rozwijać i badać ich przydatność w tym dziale nauki oraz wdrażać najlepsze rozwiązania do praktyki farmakoekonomicznej, jako że modele neuronowe mogą stanowić cenne narzędzie porównawcze i uzupełniające dla modeli klasycznych.
3. Zastosowanie modeli neuronowych umożliwia ukierunkowanie drogi farmakoterapii również pod kątem farmakoekonomicznym.
4. Sztuczne sieci neuronowe oparte o logikę rozmytą umożliwiają sprostanie obecnie obowiązującym wytycznym modelowania farmakoekonomicznego dzięki możliwości ich analizy.
5. Sztuczne sieci neuronowe jako narzędzia uniwersalne mogą być zastosowane zarówno do tworzenia modeli farmakoekonomicznych jak i epidemiologicznych, przy zastosowaniu analogicznej metodyki.
6. Jakość modelu neuronowego zbudowanego w oparciu o informacje epidemiologiczne, w dużym stopniu zależy od jakości danych, podobnie jak to jest w przypadku modeli klasycznych.
7. Istnieje możliwość stworzenia i praktycznego wykorzystania neuronowych modeli skriningowych dla chorób przewlekłych pracujących w oparciu o metody inteligencji obliczeniowej.
8. Dzięki zastosowaniu metod programowania na potrzeby Internetu istnieje możliwość udostępniania modeli neuronowych o charakterze skriningowym.



## 9 Streszczenie

Praca dotyczyła oszacowania możliwości zastosowania sztucznych sieci neuronowych w farmakoekonomice i epidemiologii. Ocenę przeprowadzono w oparciu o przykłady konkretnych rozwiązań z zakresu obu nauk. Pierwsza grupa obejmuje modelowanie efektu zastosowania środka leczniczego na potrzeby analizy farmakoekonomicznej. W przypadku zagadnień epidemiologicznych przygotowano neuronowe modele skriningowe dla wybranych chorób przewlekłych.

Ze względu na szeroki zakres podjętych prac w trakcie ich trwania wykorzystano różnorodne oprogramowanie. Podstawowym programem był symulator sieci neuronowych Nets2004, oprócz tego korzystano ze składników pakietów analizy danych oraz oprogramowania specjalizowanego, skonstruowanego specjalnie na potrzeby podjętych doświadczeń. Jego zastosowanie okazało się konieczne ze względu na rozmiar i specyficzne kodowanie danych pierwotnych, wykorzystanych w trakcie doświadczeń.

Wyniki badań potwierdzają słuszność postawionych założeń. Wykazano możliwość zastosowania modeli bazujących na inteligencji obliczeniowej jako narzędzi modelowania efektu w analizie farmakoekonomicznej. Korzystając z ograniczonej bazy wiedzy (100 przypadków włączonych do etapu doświadczalnego) udało się uzyskać 84% skuteczność systemu predykcyjnego czasu przeżycia pacjenta. Wykorzystując szybkość działania sieci neuronowych pozwala to na konstruowanie w czasie rzeczywistym modeli przeżycia dla zmieniających się warunków reżimu farmakoterapii. Wykonane doświadczenia intuicyjne działanie wytrenowanego systemu neuronowego, zgodne z obecną wiedzą z zakresu farmakoterapii chorób nowotworowych. Uzyskane wyniki mogą być w bezpośredni sposób wykorzystane do obliczania współczynników farmakoekonomicznych. Po ocenie zmieniających się kosztów terapii uzyskiwane są bezpośrednio współczynniki koszt efektywność dla porównywanych schematów leczenia. Wykazano, że metody empiryczne

jakimi są sztuczne sieci neuronowe mogą być wykorzystane jako narzędzie modelowania w farmakoekonomice mimo ograniczenia wynikającego z charakteru „czarnej skrzynki”. Potwierdzają to wyniki porównania z klasycznymi modelami statystycznymi opartymi o regresję logistyczną. Wyniki testów generalizacji wskazują na przewagę modeli neuronowych.

Dowiedziano również iż sieci neuronowe mogą stanowić mechanizm wnioskowania dla systemu ekspertowego określającego ryzyko zapadalności na choroby przewlekłe. W przypadku systemów skriningowych bazujących na danych epidemiologicznych podkreślenia wymagają fakty możliwości wykorzystania prostych danych o charakterze epidemiologicznym oraz możliwość szerokiego udostępnienia działającego narzędzia dzięki technologii internetowej. Jako, że celem pracy było m.in. wykazanie możliwości zaimplementowania podobnych narzędzi do codziennej pracy farmaceuty lub lekarza, korzystając z technologii Java udostępniono działającą sieć neuronową za pośrednictwem internetu. Narzędzie jest dostępne pod adresem <http://www.cyf-kr.edu.pl/~mfpolak/>.

Wyniki badań potwierdzają słuszność postawionych założeń. Uniwersalny charakter wykorzystanych technik modelowania neuronowego pozwala na skuteczne zastosowanie narzędzi z zakresu inteligencji obliczeniowej w farmakoekonomice i epidemiologii.

## 10 Piśmiennictwo

1. 2003 European Society of Hypertension–European Society of Cardiology guidelines for the management of arterial hypertension. *Journal of Hypertension*. 2003, 21: 1011–1053.
2. ABPI-Government Strategic Working Group. *Guidance on Good Practice in the Conduct of Economic Evaluations of Medicines*. London: Association of the British Pharmaceutical Industry; 1994.
3. Alaskerov E., Freisleben B., Rao B. CARDWATCH: A Neural Network Based Data Mining System for Credit Card Fraud Detection, *Proceedings of the IEEE/IAFE*, 1997: 220-226.
4. Albert U.S., Schulz K.D. Short Version of the Guideline: Early Detection of Breast Cancer in Germany An evidence-, consensus-, and outcome-based guideline according to the German Association of the Scientific Medical Societies (AWMF) and the German Agency for Quality in Medicine (AeZQ). *J Cancer Res Clin Oncol*. 2004, 130: 527–536.
5. Aminzadeha F., Barhen J., Glover C.W., Toomarian N.B. Reservoir parameter estimation using a hybrid neural network. *Computers & Geosciences* 26, 2000: 869-875.
6. Anagnostou T., Remzi M., Djavan B. Artificial Neural Networks for Decision-Making in Urologic Oncology. *Rev Urol*. 2003, 5(1): 15–21.
7. Anderson K.M. et al. Cardiovascular risk profiles. *Am Heart J*, 1991; 121: 293-298.
8. Annemans L., Moeremans K., Mutimer D., Schneeberger H., Milligan D., Kubin M. Modeling Cost and Cost-Effectiveness of Different AMV Management Startegies in Liver Transplant Recipients as a Support for Current and Future Decision Making. *Value in Health* 2002; 5(4): 347-358.
9. Antowiak M., Chałasińska-Macukow K. Fingerprint identification by using artificial neural network with optical wavelet preprocessing. *Opto-Electronics Review*, 2003; 11(4): 327–337.

10. Badiru A.B., Sieger D.B. Neural network as a simulation metamodel in economic analysis of risky projects. *European Journal of Operational Research* 1998, 105: 130-142.
11. Baesens B., Viaene S., Van den Poel D., Vanthienen J., Dedene G. Bayesian neural network learning for repeat purchase modelling in direct marketing. *European Journal of Operational Research* 2002, 138(1): 191-211.
12. Bailey N.T.J. *The Mathematical Theory of Infectious Diseases*. New York, Hafner Press 1975.
13. Basskin L.E., Performing a pharmacoeconomic analysis of a disease management pathway or guideline. *Formulary* 2003; 38: 547-551.
14. Batuello J.T., Crawford E.D. Artificial neural networks in prostate cancer. *LabMedica International* 2001; 18(3): 10-11.
15. Bautista D., Arana E., Martí-Bonmatí L., Paredes R. Validation of Logistic Regression Models in Small Samples: Application to Calvarial Lesions Diagnosis. *J Clin Epidemiol.* 1999 52(3): 237–241.
16. Bootman J.L., Townsend R.J., McGhan W.F., *Principles of Pharmacoeconomics*. Harvey Whitney Books Company, Cincinnati 1991.
17. Bradley A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition.* 1997, 30(7): 1145-1159.
18. Brennan A., Akehurst R. Modelling in Health Economic Evaluation What is its Place? What is its Value? *Pharmacoeconomics* 2000; 17 (5): 445-459.
19. Bubern H., Hillig H. Artificial Neural Network and Fuzzy Clustering - New Tools for Evaluation of Depth Profile Data? *Mikrochim. Acta* 2000, 133: 95-103.
20. Burnier M., Schneider M.P., Chioléro A., Fallab C.L., Brunner H.R., Objective monitoring of drug compliance: an important step in the management of hypertension resistant to drug therapy. *American Journal of Hypertension.* 1999, 12(4): 129.

21. Byers A.L., Allore H., Gill T.M., Peduzzi P.N. Application of negative binomial modeling for discrete outcomes. A case study in aging research. *J Clin Epidemiol.* 2003, 56: 559–564.
22. Cacciafesta M. et al. Neural network analysis in predicting 2-year survival in elderly people: a new statistical–mathematical approach. *Archives of Gerontology and Geriatrics* 2001, 32: 35–44.
23. Canadian Coordinating Office for Health Technology Assessment. Guidelines for economic evaluation of pharmaceuticals: Canada. 2nd ed. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA); 1997.
24. Casey B.M., Lucas M.J., McIntire D.D., Leveno K.J. Pregnancy outcomes in women with gestational diabetes compared with the general obstetric population. *Obstetrics and Gynecology.* 1997, 90: 869–873.
25. Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Data. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, [1999]. <http://www.cdc.gov/brfss/>
26. Chaw Koh Yong, Choo Min Lim, Plumbley, M. Beighton, D.; Davidson, R. Identification of dental bacteria using statistical and neural approaches. *Neural Information Processing, 2002. ICONIP'02 Proceedings of the 9th International Conference 18-22 November 2002, 2002, 2: 606 – 610.*
27. Chen L., Chen M.H., Larson M.G., Evans J., Benjamin E.J., Levy D. Risk Factors for Syncope in a Community-Based Sample (The Framingham Heart Study). *American Journal of Cardiology* 2000; 85: 1189–1193.
28. Chuang J.-H. et al. A Web-Based System For Prediction Of Coronary Heart Disease Risk Using The Framingham Algorithm. *The American Medical Informatics Association 2000 Symposium, 2000;* <http://www.amia.org/meetings/archive/f3.html>.

29. Cichosz, P. Systemy uczące się. Wydawnictwo Naukowo-Techniczne, Warszawa 2000.
30. Çoban H. Application of an Artificial Neural Network (ANN) for the Identification of Grapevine (*Vitis vinifera* L.) Genotypes. *Asian Journal of Plant Sciences* 2004, 3 (3): 340-343.
31. Crawford E.D., et al. Use of artificial intelligence technology to predict lymph node spread in men with clinically localized prostate cancer. *Cancer*. 2000; 88(9):2105-2109.
32. Dancourta V., Quantina C., Abrahamowicz M., Bingueta C., Alioum A., Faivred J. Modeling recurrence in colorectal cancer. *J Clin Epidemiol*. 2004, 57: 243–251.
33. Davies R. Roderick P., Raftery J. The evaluation of disease prevention and treatment using simulation models. *European Journal of Operational Research* 2003, 150: 53–66.
34. Dąbrowska B. Nowa definicja zawału serca. *Medycyna Praktyczna* 2001/03.
35. Delen D., Walker G., Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. *Artificial Intelligence in Medicine* 2005, 34: 113–127.
36. Djavan B., Remzi M., Zlotta A., Seitz C., Snow P., Marberger M. Novel Artificial Neural Network for Early Detection of Prostate Cancer. *Journal of Clinical Oncology*. 2002, 20(4): 921-929.
37. Drożdż J. et al. Artificial Neural Network Predicts Mortality in Patients With Acute Myocardial Infarction. *Polish Multicenter Web-Based Trial on Acute Myocardial Infarction: POL-WEB-AMI. JACC* 2001, 2: 494A.
38. Drummond M.F., O'Brien B., Stoddart G., Torrance G.W. *Metody badań ekonomicznych programów ochrony zdrowia (tłum. z języka angielskiego pod red. J. Splawińskiego)*. ViaMedica, Gdańsk 2003.
39. Duch W., Korbicz J., Rutkowski L., Tadeusiewicz R. – red. *Biocybernetyka i Inżynieria Biomedyczna. Tom 6 Sieci Neuronowe*. Akademicka Oficyna Wydawnicza EXIT, Warszawa 2000.

40. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. *J. Am. Stat. Assoc.*, 1983; 78: 316–331.
41. Elliott W.J. Medication non-compliance is a risk factor for beginning dialysis in a tertiary hypertension clinic. *American Journal of Hypertension*. 1995, 8(4): 99A.
42. Faddy M.J., Taylor G.J., Stochastic Modelling of the Onset of *Brionchiolitis Obliterans Syndrome* Following Lung Transplantation: An Analysis of Risk Factors. *Mathematical and Computer Modelling* 2003; 38: 1185-1189.
43. Feuer E.J., Etzioni R., Cronin K.A., Mariotto A. The use of modeling to understand the impact of screening on US mortality: examples from mammography and PSA testing. *Statistical Methods in Medical Research* 2004, 13: 421-442.
44. Frauenthal J. C., *Mathematical Modeling in Epidemiology*. Springer-Verlag, New York 1980.
45. Gamito E.J., Stone N.N., Batuello J.T., Crawford E.D. Use of artificial neural networks in the clinical staging of prostate cancer: implications for prostate brachytherapy. *Techniques in Urology*. 2000; 6(2): 60-63.
46. Gansky S.A. Dental Data Mining: Potential Pitfalls and Practical Issues. *Adv Dent Res* 2003, 17: 109-114.
47. Goh W.Y., Lim C.P., Peh K.K., Predicting Drug Dissolution Profiles With an Ensemble of Boosted Neural Networks: A Time Series Approach. *IEEE T. Neur. Networ.* 2003, 11(2): 459-463.
48. Gold M.R., Siegel J.E., Russell L.B., Weinstein M.C., *Cost-Effectiveness in Health and Medicine*. Oxford University Press, Oxford 1996.
49. Goldstein R.Z. et al. Modeling Preclinical Cardiovascular Risk for Use in Epidemiologic Studies. *Am J Epidemiol* 2001, 154: 765–76.
50. Goralwalla I.A., Ozsu M.T., Szfron D. Modeling Medical Trials in Pharmacoeconomics using Temporal Object Model. *Computers in*

- Biology and Medicine Volume: 27, Issue: 5, September, 1997, pp. 369-387.
51. Greenberg R. A., Sacks D. B. Screening for diabetes: is it warranted? *Clinica Chimica Acta*. 2002, 315: 61–69.
  52. Gyrd-Hansen D. The relative economics of screening for colorectal cancer, breast cancer and cervical cancer. *Critical Reviews in Oncology Hematology* 1999; 32: 133–144.
  53. Hammad T.A., Abdel-Wahab M.F., DeClaris N., El-Sahly A., El-Kady N., Strickland G.T. Comparative evaluation of the use of artificial neural networks for modelling the epidemiology of schistosomiasis mansoni. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 1996, 90: 372-376.
  54. Hashem S. Sensitivity Analysis of Feedforward Artificial Neural Networks with Differentiable Activation Functions. *Proceedings of the 1992 International Joint Conference on Neural Networks, Baltimore, IEEE, New Jersey, 1992: 419-424.*
  55. Hertz J., Krogh A., Palmer R.G. *Wstęp do teorii obliczeń neuronowych*. Wydawnictwa Naukowo-Techniczne, Warszawa 1995.
  56. Hethcote H.W. Mathematics of infectious diseases. *SIAM Review* 2000;42:599-653.
  57. Hosmer D.W., Lemeshow S. *Applied Logistic Regression*. John Wiley & Sons, New York 1989.
  58. <http://bcra.nci.nih.gov/brc/>
  59. <http://citeseer.ist.psu.edu/meyers01applying.html>
  60. <http://hin.nhlbi.nih.gov/atpiii/calculator.asp?usertype=pub#sbp>
  61. <http://www.cdc.gov>
  62. <http://www.cukrzyca.info.pl>
  63. <http://www.nci.nih.gov>
  64. <http://www.nhlbi.nih.gov/about/framingham/index.html>
  65. <http://www.prostatecalculator.org>
  66. [http://www.who.int/cardiovascular\\_diseases/resources/atlas/en/](http://www.who.int/cardiovascular_diseases/resources/atlas/en/)



67. [http://www.who.int/mip/2003/other\\_documents/en/globalburdenofdisease.pdf](http://www.who.int/mip/2003/other_documents/en/globalburdenofdisease.pdf)
68. <http://www.yourdiseaserisk.harvard.edu/>
69. Jang J.-S.R., Sun C.-T. Neuro-Fuzzy Modeling and Control. Proceedings of the IEEE, 1995; 83(3): 378-406.
70. Jones J., Wilson A., Parker H., Wynn A., Jagger C., Spiers N., Parker G. Economic evaluation of hospital at home versus hospital care: cost minimisation analysis of data from randomised controlled trial. BMJ 1999; 319: 1547-1550.
71. Khine M.L., Winklestein A., Copel J.A. Selective screening for gestational diabetes mellitus in adolescent pregnancies. Obstetrics and Gynecology. 1999, 93: 738-742.
72. Korbicz J., Obuchowski A., Uciński D. Sztuczne sieci neuronowe. Podstawy i zastosowania. Akademicka Oficyna Wydawnicza PLJ, Warszawa 1994.
73. Kosmidis P. Chemotherapy in NSCLC: historical review. Lung Cancer 2002; 38: S19-S22.
74. Lafuente R., et al. Design and test of neural networks and statistical classifiers in computer-aided movement analysis: a case study on gait analysis. Clinical Biomechanics 1997; 13(3): 216-229.
75. Lapuerta P., Rajan S., Bonacini M. Neural Networks As Predictors of Outcomes in Alcoholic Patients With Severe Liver Disease. Hepatology 1997, 25(2): 302-306.
76. Last J.M. A dictionary of epidemiology, 4th ed., Oxford: Oxford University Press.
77. Lee C.W., Park J.-A. Assessment of HIV/AIDS-related health performance using an artificial neural network. Information & Management 2001; 38: 231-238.
78. Liew D., McNeil J.J, Peeters A., Lim S.S, Vos T. Epidemiological modelling (including economic modelling) and its role in preventive drug therapy. MJA, 2002; 177: 364-367.

79. Lumley T., Kronmal R.A., Cushman M., Manolio T.A., Goldstein S. A stroke prediction score in the elderly: validation and Web-based application. *Journal of Clinical Epidemiology* 2002; 55: 129–136.
80. Luxhoj J.T., Williams T.P. Integrated decision support for aviation safety inspectors. *Finite Elements in Analysis and Design* 1996, 23: 381-403.
81. Lynch J.H. et al. Therapeutic Strategies for Localized Prostate Cancer. *Rev Urol.* 2001, 3(suppl 2): S39–S48.
82. Marble R.P., Healy J.C. A neural network approach to the diagnosis of morbidity outcomes in trauma care. *Artificial Intelligence in Medicine* 1999, 15: 299–307.
83. Martin J.K., Hirschberg D.S. Small Sample Statistics for Classification Error Rates I: Error Rate Measurements. <http://citeseer.ist.psu.edu/148875.html>
84. Marx G., Harper P. Non-platinum gemcitabine combinations in non-small cell lung cancer. *Lung Cancer* 2002; 38: S51-S54.
85. May R., Anderson R. *Infectious Diseases of Humans: Dynamics and Control.* Oxford, Oxford University Press 1992.
86. Mendyk A. Sztuczne sieci neuronowe jako uniwersalne narzędzia do modelowania w technologii postaci leku i biofarmacji. Praca doktorska, Kraków 2004.
87. Mendyk A., Jachowicz R. Neural network as a decision support system in the development of pharmaceutical formulation—focus on solid dispersions. *Expert Systems with Applications* 2005, 28(2): , 285-294.
88. Minsky M., Papert S. *Perceptrons.* MIT Press, Cambridge 1969.
89. Mohamed E.I., Maiolo C., Linder R., Pöppel S.J., De Lorenzo A. Artificial neural network analysis: a novel application for predicting site-specific bone mineral density. *Acta Diabetol*, 2003; 40: S19–S22.

90. Montano J.J., Palmer A. Numeric sensitivity analysis applied to feedforward neural networks. *Neural Comput & Applic*, 2003; 12: 119–125.
91. Mullins C.D., Ogilvie S. Emerging Standardization in Pharmacoeconomics. *Clinical Therapeutics* 1998; 20(6): 1194-1202.
92. Nuijten M.J.C., Rutten F. Combining a Budgetary-Impact Analysis and a Cost-Effectiveness Analysis Using Decision-Analytic Modelling Techniques. *Pharmacoeconomics* 2002; 20 (12): 855-867.
93. Ohno-Machado L. Modeling Medical Prognosis: Survival Analysis Techniques. *Journal of Biomedical Informatics* 2001; 34: 428-439
94. Ontario Guidelines for Economic Analysis of Pharmaceutical Products. Ontario Ministry of Health and Long-Term Care, 1994.
95. Orlewska E. Rola i wartość modelowania w ocenie ekonomicznej programów zdrowotnych. *Farmakoeconomika* 2002; 2
96. Orlewska E., *Podstawy farmakoeconomiki*. Unimed, Warszawa 1999.
97. Ortiz J. et al. One-Year Mortality Prognosis in Heart Failure: A Neural Network Approach Based on Echocardiographic Data. *JACC* 1995, 26(7): 1586-1593.
98. Osowski S. *Sieci neuronowe w ujęciu algorytmicznym*. Wydawnictwo Naukowo-Techniczne, Warszawa 1996.
99. Ottenbacher K.J., Smith P.M., Illig S.B., Linn R.T., Fiedler R.C., Granger C.V. Comparison of logistic regression and neural networks to predict rehospitalization in patients with stroke. *Journal of Clinical Epidemiology* 2003; 54: 1159–1165.
100. Paesmans M., Response to Chemotherapy has Predictive Value for Further Survival of Patients with Advanced Non-small Cell Lung Cancer: 10 Years Experience of the European Lung Cancer Working Party. *European Journal of Cancer* 1997; 33(14): 2326-2332.
101. Polak S., Mendyk A. Artificial Intelligence Technology as a Tool for Initial GDM Screening. *Expert Systems with Applications*. 2004, 26(4): 455-460.

102. Polak S., Mendyk A., Artificial Neural Networks as an Engine of Internet Based Hypertension Prediction Tool in Medical and Care Compunetics part 1 – Studies in Health Technology and Informatics, IOS Press, Amsterdam, The Netherlands, 2004.
103. Polskie wytyczne przeprowadzania badań farmakoekonomicznych (projekt), Polskie Towarzystwo Farmakoekonomiczne (<http://www.farmakoekonomika.pl/wytyczne.htm>)
104. Potter S.T. et al. Genetically Engineered Neural Networks for Predicting Prostate Cancer Progression After Radical Prostatectomy. Urology. 1999, 54: 791–795.
105. Principles of Epidemiology. An Introduction to Applied Epidemiology and Biostatistics. U.S. Department of Health and Human Services, Public Health Service Centers for Disease Control and Prevention (CDC).
106. Puska P. Successful prevention of non-communicable diseases: 25 year experiences with North Karelia Project in Finland. Public Health Medicine 2002, 4(1): 5-7.
107. Racheta B., Siemiatycki J., Abrahamowicz M., Leffondre K. A flexible modeling approach to estimating the component effects of smoking behavior on lung cancer. J Clin Epidemiol. 2004, 57: 1076–1085.
108. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus (2003). Diabetes Care. 2003, 26(Supp. 1): 5–20.
109. Ripley R.M. Neural Network Models for Breast Cancer Prognosis. PhD thesis Department of Engineering Science, University of Oxford.
110. Rippin J.D., Barnett A.H., Bain S.C. Cost-Effective Strategies in the Prevention of Diabetic Nephropathy. Pharmacoeconomics 2004; 22(1): 9-28.
111. Robinson R. Costs and cost-minimisation analysis. BMJ 1993; 307: 726–728.
112. Rutkowska D., Piliński M., Rutkowski L. Sieci neuronowe, algorytmy genetyczne i systemy rozmyte. PWN, Warszawa-Łódź 1997.

113. Santos-Garcia G., Varela G., Novoa N., Jimenez M.F. Prediction of postoperative morbidity after lung resection using an artificial neural network ensemble. *Artificial Intelligence in Medicine* 2004, 30: 61–69.
114. Saxena P., McDonald R., Gull S., Hyder N. Diagnostic scanning for suspected scaphoid fractures: an economic evaluation based on cost-minimisation models. *Injury, Int. J. Care Injured* 2003; 34: 503–511.
115. Scalon J.D., Vidal Melo M.F., Panerai R.B. A method for predicting the utilization of medical technology in neonatal care. *Technology and Health Care* 1996; 3: 273-278.
116. Schiller J., Tilden D., Aristides M., Lees M., Kielhorn A., Maniadakis N., Bhallad S. Retrospective cost analysis of gemcitabine in combination with cisplatin in non-small cell lung cancer compared to other combination therapies in Europe. *Lung Cancer* 2004; 43: 101–112.
117. Schneider G., Schrödl W., Wallukat G., Müller J., NISSEN E., Röspeck W., Wrede P., Kunze R. Peptide design by artificial neural networks and computer-based evolutionary search. *Proc. Natl. Acad. Sci. USA* 1998, 95: 12179–12184.
118. Seventh Report of the Joint National Committee on the Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC 7). U. S. DEPARTMENT OF HEALTH AND HUMAN SERVICES; <http://www.nhlbi.nih.gov>
119. Shepperd S., Harwood D., Gray A., Vessey M., Morgan P. Randomised controlled trial comparing hospital at home care with inpatient hospital care. II: cost minimisation analysis. *BMJ* 1998; 316: 1791 - 1796.
120. Skowron A. Analiza farmakoekonomiczna terapii niedrobnokomórkowego raka płuc. Praca doktorska. Kraków 2003.
121. Stanisław A. Przystępny kurs statystyki. StatSoft Polska, Kraków 2000.

122. Strolka I., Accardo A., Dreossi D., Vittur F., Toffanin R., Frollo I. Segmentation of Magnetic Resonance Microimages of Trabecular Bone: Classifiers and Markov Random Field Model. *Measurement Science Review*. 2003, 3(2): 79-82.
123. Tadeusiewicz R. *Sieci Neuronowe*. Akademicka Oficyna Wydawnicza RM, Warszawa 1993.
124. Tandon R., Adak S., Kaye J.A. Neural Networks for Longitudinal Studies in Alzheimer's Disease. *GE Global Research* 2003; GRC 294.
125. Thomas M., et al. Development of treatment strategies in locally advanced non-small cell lung cancer. *Lung Cancer* 2001; 33 Suppl.1: S91-S97.
126. Thongprasert S., Sanguanmitra P., Juthapan W., Clinch J. Relationship between quality of life and clinical outcomes in advanced non-small cell lung cancer: best supportive care (BSC) versus BSC plus chemotherapy. *Lung Cancer* 1999; 24: 17-24.
127. Tjan-Heijnen V.C.G., Groen H.J.M., Schramel F.M.N.H., Stoter G. Consensus conference on palliative treatment of stage IV non-small cell lung cancer. Consensus meeting of the Netherlands Society for Medical Oncology and the Taskforce on Lung Cancer of the Dutch Society of Physicians for Lung Diseases and Tuberculosis. *The Netherlands Journal of Medicine* 2001;58: 52-61.
128. Tu J.V., Austin P.C., Walld R., Roos L., Agras J., McDonald K.M. Development and Validation of the Ontario Acute Myocardial Infarction Mortality Prediction Rules. *Journal of the American College of Cardiology* 2001; 37(4): 992-997.
129. Turner J.V., Maddalena D.J., Cutler D.J. Pharmacokinetic parameter prediction from drug structure using artificial neural networks. *International Journal of Pharmaceutics* 2004, 270: 209-219.
130. Veenstra D.L., Higashi M.K., Phillips K.A. Assessing the Cost-Effectiveness of Pharmacogenomics. *AAPS Pharmsci* 2000; 2(3) (<http://www.pharmsci.org/>).

131. Walczak S., Pofahl W.E., Scorpio R.J. A decision support tool for allocating hospital bed resources and determining required acuity of care. *Decision Support Systems* 2002, 34: 445–456.
132. Walczak S., Scharf J.E. Reducing surgical patient costs through use of an artificial neural network to predict transfusion requirements. *Decision Support Systems* 2000, 30: 125–138.
133. Wang T.J., Massaro J.M., Levy D., Vasan R.S., Wolf P.A., D'Agostino R.B., Larson M.G., Kannel W.B., Benjamin E.J. A Risk Score for Predicting Stroke or Death for Individuals with New-Onset Atrial Fibrillation in the Community: The Framingham Heart Study. *JAMA* 2003;290:1049-1056.
134. Wei J.T. Understanding Artificial Neural Networks and Exploring Their Potential Applications for the Practicing Urologist. *Urology*. 1998, 52: 161–172.
135. Weinstein M.C., O'Brien B., Hornberger J., Jackson J., Johannesson M., McCabe C., Luce B.R. Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices—Modeling Studies. *Value in Health* 2003; 6(1): 9-17.
136. Weitkunat R., Wildner M. Exploratory causal modeling in epidemiology: Are all factors created equal? *J Clin Epidemiol*. 2002, 55: 436–444.
137. Werbos P.J. Beyond regression: new tool for prediction and analysis in the behavioral sciences. PhD thesis, Harvard University, Cambridge, MA 1974.
138. Whitley D. Genetic Algorithms and Neural Networks. in *Genetic Algorithms in Engineering and Computer Science*. Periaux J., Winter G. – eds. New York, John Wiley & Sons Ltd. 1995.
139. Wilson P.W. F., An Epidemiologic Perspective of Systemic Hypertension, Ischemic Heart Disease, and Heart Failure. *Am J Cardiol*. 1997, 80(9B): 3J–8J.

140. Wong N.D. Preventing heart disease by controlling hypertension: Impact of hypertensive subtype, stage, age, and sex. *American Heart Journal*. 2003, 145(5): 888-895.
141. Yager R.R., Filev D.P. *Podstawy modelowania i sterowania rozmytego*. Wydawnictwo Naukowo-Techniczne, Warszawa 1995.
142. Zadeh L.A. Fuzzy Sets. *Information and Control*, 1965; 8: 338-353.
143. Zhang G., Hu M.Y., Patuwo B.E., Indro D.C. Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. *European Journal of Operational Research* 1999, 116: 16-32.



## 11 Spis równań

Równanie 1 Wzór funkcji progowej (jednostkowej).....	12
Równanie 2 Wzór funkcji sigmoidalnej unipolarnej.....	12
Równanie 3 Wzór funkcji tangens hiperboliczny. ....	12
Równanie 4 Postać funkcji błędu w uczeniu z nauczycielem. ....	17
Równanie 5 Postać jednej z funkcji celu – błędu średniokwadratowego. .	17
Równanie 6 Matematyczny zapis zmian wartości wektora wag w zależności od wartości błędu na wyjściu sieci. ....	19
Równanie 7 Schemat modyfikacji wag w kolejnych krokach uczenia .....	19
Równanie 8 Technika momentu. ....	22
Równanie 9 Postać funkcji Karayannisa.....	23
Równanie 10 Proponowana postać funkcji błędu $\varphi$ w algorytmie Karayannisa.....	24
Równanie 11 Alternatywna funkcja celu – funkcja logarytmiczna.....	24
Równanie 12 Funkcja Hampela. ....	24
Równanie 13 Funkcja celu zawierająca funkcję Hampela .....	25
Równanie 14 Reguła Hebba. ....	29
Równanie 15 Funkcja gaussowska.....	34
Równanie 16 Współczynnik koszt-efektywność (CER).....	50
Równanie 17 Inkrementalny współczynnik koszt-efektywność (ICER). ...	50
Równanie 18 Schemat obliczenia wartości preferencji metodą handlowania czasem w przypadku chorób przewlekłych.....	55
Równanie 19 Schemat obliczenia wartości preferencji metodą handlowania czasem w przypadku chorób przejściowych. ....	56
Równanie 20 Schemat obliczenia wartości preferencji metodą loterii w przypadku chorób przewlekłych.....	57
Równanie 21 Schemat obliczenia wartości preferencji metodą loterii w przypadku chorób przejściowych. ....	58
Równanie 22 Postać funkcji regresji logistycznej. ....	76
Równanie 23 Zapis matematyczny logistycznego modelu regresji.....	76
Równanie 24 Wzór umożliwiający obliczenie błędu średniokwadratowego (RMSE). ....	85

## 12 Spis schematów

Schemat 1 Model sztucznego neuronu wzorowany na neuronie biologicznym.....	10
Schemat 2 Wykresy funkcji progowej (jednostkowej), liniowej i logistycznej (sigmoida unipolarna).....	13
Schemat 3 Przebieg funkcji logistycznej w zależności od parametru $\beta$ ....	13
Schemat 4 Schemat sztucznej sieci neuronowej dwuwarstwowej.....	14
Schemat 5 Schemat sieci neuronowej rekurencyjnej.....	15
Schemat 6 Schemat sieci neuronowej komórkowej.....	16
Schemat 7 Schemat blokowy uczenia nadzorowanego sieci. ....	18
Schemat 8 Pierwszy etap uczenia sieci wielowarstwowej metodą wstecznej propagacji błędów.....	20
Schemat 9 Drugi etap uczenia sieci wielowarstwowej metodą wstecznej propagacji błędów.....	20
Schemat 10 Sieć uczona ze wzmocnieniem (z krytykiem).....	27
Schemat 11 Algorytm Kohonena .....	31
Schemat 12 Rozkład funkcji przynależności dla systemu rozmytego (a) oraz ostrego (b) .....	34
Schemat 13 System rozmyty z blokiem rozmywania (fuzyfikator) oraz wyostrzania (defuzyfikator). ....	35
Schemat 14 Graficzne odwzorowanie uczenia sieci neuronowej rozmytej. ....	35
Schemat 15 Zależność między preferencją oraz wartością i użytecznością. ....	53
Schemat 16 Graficzny opis metody handlowania czasem dla chorób przewlekłych [38,96]. ....	55
Schemat 17 Graficzny opis metody handlowania czasem dla chorób przejściowych. ....	56
Schemat 18 Ocena preferencji dla stanu przewlekłego przy zastosowaniu metody loterii. ....	57
Schemat 19 Ocena preferencji dla stanu przejściowego przy zastosowaniu metody loterii. ....	58
Schemat 20 Wykres funkcji logistycznej.....	76
Schemat 21 Sposób opisu testowanych architektur sieci neuronowych... 86	
Schemat 22 Zakres modyfikacji danych pierwotnych - diagram przebiegu doświadczeń. ....	92
Schemat 23 Etap I - modelowanie z wykorzystaniem danych pierwotnych .....	93
Schemat 24 Wykres zależności wartości rzeczywistej do przewidzianej przez model dla architektury 2hid M-SLIDE, uczonej na danych pierwotnych z wyjściem ciągłym.....	94
Schemat 25 Porównanie wartości rzeczywistych i przewidzianych przez model dla architektury 2hid M-SLIDE, uczonej na danych pierwotnych z wyjściem ciągłym; analiza trendów. ....	94

Schemat 26 Wykres zależności wartości rzeczywistej do przewidzianej przez model dla architektury 10_7_5_2hid sigma, uczonej na danych pierwotnych z wyjściem ciągłym.....	95
Schemat 27 Porównanie wartości rzeczywistych i przewidzianych przez model dla architektury 10_7_5_2hid sigma, uczonej na danych pierwotnych z wyjściem ciągłym; analiza trendów. ....	95
Schemat 28 Etap II – wstępne formatowanie danych oraz modelowanie z wykorzystaniem danych przeformatowanych – zaszumianie danych. 96	
Schemat 29 Etap III – wstępne formatowanie danych oraz modelowanie z wykorzystaniem danych przeformatowanych – dyskretyzacja danych. ....	100
Schemat 30 Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na danych o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść. ....	104
Schemat 31 Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na danych o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść z funkcją Hampela jako funkcją kosztu. ....	106
Schemat 32 Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na danych o binarnym charakterze wyjścia (30 tygodni) dla 30 wejść. ....	107
Schemat 33 Etap IV – Modelowanie z wykorzystaniem danych przeformatowanych – wyjście binarne, 30 wejść, zaszumianie danych. ....	109
Schemat 34 Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 2x2%) o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść. ....	111
Schemat 35 Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 2x5%) o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść. ....	112
Schemat 36 Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 2x10%) o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść. ....	114
Schemat 37 Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 5x5%) o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść. ....	115
Schemat 38 Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 10x5%) o binarnym charakterze wyjścia (35 tygodni) dla 30 wejść. ....	117
Schemat 39 Wykresy ROC oraz wartości AUROC dla najlepszych modeli uzyskanych na zaszumionych danych (parametry szumu - 5x5%) o binarnym charakterze wyjścia (30 tygodni) dla 30 wejść. ....	118
Schemat 40 Etap V – Modelowanie z wykorzystaniem danych przeformatowanych – wyjście binarne, 31 wejść, zaszumianie danych. ....	120
Schemat 41 Wykres zależności wartości rzeczywistej do przewidzianej przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z domyślną wartością prognozy. ....	121

Schemat 42 Porównanie wartości rzeczywistych i przewidzianych przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z domyślną wartością proggu; analiza trendów.....	122
Schemat 43 Wykres zależności wartości rzeczywistej do przewidzianej przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z domyślną wartością proggu. ....	123
Schemat 44 Porównanie wartości rzeczywistych i przewidzianych przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z domyślną wartością proggu; analiza trendów.....	123
Schemat 45 Wykres zależności wartości rzeczywistej do przewidzianej przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z modyfikowaną wartością proggu. ....	124
Schemat 46 Porównanie wartości rzeczywistych i przewidzianych przez model dla architektury 2hid fsr, uczonej na danych opisanych przez 31 wejść z domyślną wartością proggu; analiza trendów.....	125
Schemat 47 Schemat zagadnień z zakresu modelowania epidemiologicznego .....	128
Schemat 48 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka cukrzycy ciężarnych uzyskane na danych pierwotnych, balansowanych dodatnio. ....	133
Schemat 49 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka cukrzycy ciężarnych uzyskane na danych przekodowanych, balansowanych dodatnio. ....	135
Schemat 50 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych pierwotnych (38 wejść), balansowanych ujemnie. ....	141
Schemat 51 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych pierwotnych (38 wejść), balansowanych dodatnio. ....	142
Schemat 52 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych pierwotnych (8 wejść), balansowanych ujemnie. ....	143
Schemat 53 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka ostrego incydentu wieńcowego uzyskane na danych pierwotnych (28 wejść), balansowanych ujemnie.....	151
Schemat 54 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka ostrego incydentu wieńcowego uzyskane na danych pierwotnych (28 wejść), balansowanych dodatnio.....	152
Schemat 55 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka ostrego incydentu wieńcowego uzyskane na danych pierwotnych (8 wejść), balansowanych ujemnie. ....	153
Schemat 56 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka ostrego incydentu wieńcowego uzyskane na danych pierwotnych (28 wejść), balansowanych dodatnio.....	154
Schemat 57 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka wystąpienia choroby nowotworowej uzyskane na danych pierwotnych (101 wejść), balansowanych ujemnie.....	157

Schemat 58 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka wystąpienia choroby nowotworowej uzyskane na danych pierwotnych (101 wejść), balansowanych dodatnio.....	158
Schemat 59 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka wystąpienia choroby nowotworowej uzyskane na danych pierwotnych (18 wejść), balansowanych ujemnie.....	159
Schemat 60 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka wystąpienia choroby nowotworowej uzyskane na danych pierwotnych (18 wejść), balansowanych dodatnio.....	160
Schemat 61 Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych przekodowanych – różne metody modelowania (cukrzyca ciężarnych, 8 wejść). ....	182
Schemat 62 Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych przekodowanych – różne metody modelowania (nadciśnienie tętnicze, 6 wejść). ....	185
Schemat 63 Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych pierwotnych – różne metody modelowania (ostry incydent wieńcowy, 28 wejść). ....	188
Schemat 64 Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych pierwotnych – różne metody modelowania (choroba nowotworowa, 101 wejść).....	190

### 13 Spis wykresów

Wykres 1 Przebieg zmian odsetka pacjentów w zależności od stopnia redukcji dawki leków.....	164
Wykres 2 Zamodelowany przebieg zmian okresu przeżycia pacjentów w zależności od charakteru środka p/wymiotnego. ....	171
Wykres 3 Zamodelowany przebieg zmian okresu przeżycia pacjentów w zależności od charakteru środka p/wymiotnego. ....	172

## 14 Spis tabel

Tabela 1 Charakterystyka typów analiz farmakoekonomicznych oraz analizy kosztów.....	44
Tabela 2 Metody pomiaru preferencji [38].....	53
Tabela 3 Konfiguracja stacji roboczych wykorzystywanych w trakcie prac. .....	80
Tabela 4 Wykaz szpitali, z których pochodzą dane wykorzystane w trakcie badań. ....	82
Tabela 5. Zmienne opisujące rekordy pacjentów włączonych do badania (niedrobnokomórkowy rak płuc, 30 wejść).....	83
Tabela 6 Opis statystyczny zmiennych wejściowych.....	84
Tabela 7 Opis statystyczny wartości modelowanej (przeżycia). ....	85
Tabela 8 Zakres zaszumiania danych pierwotnych.....	87
Tabela 9 Liczebność poszczególnych klas przy podziale binarnym z zastosowaniem prognozy 30 i 35 tygodni przeżycia. ....	88
Tabela 10 Liczebność poszczególnych klas przy podziale na cztery klasy (długość pierwszego przedziału – 15 tyg.). ....	90
Tabela 11 Liczebność poszczególnych klas przy podziale na cztery klasy (długość pierwszego przedziału – 35 tyg.). ....	90
Tabela 12 Liczebność poszczególnych klas przy podziale na trzy klasy (długość pierwszego przedziału – 15 tyg.). ....	90
Tabela 13 Liczebność poszczególnych klas przy podziale na trzy klasy (długość pierwszego przedziału – 35 tyg.). ....	90
Tabela 14 Architektury testowane w trakcie doświadczeń na danych pierwotnych – najlepsze uzyskane wyniki. ....	93
Tabela 15 Architektury testowane w trakcie doświadczeń na danych pierwotnych z zastosowaniem funkcji Hampela – najlepsze uzyskane wyniki .....	96
Tabela 16 Parametry zaszumiania danych; ilość rekordów zbiorów wynikowych – wyjście ciągłe.....	97
Tabela 17 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem ciągłym – najlepsze wyniki dla zbioru o parametrach parametrów szumu 2x2%. ....	97
Tabela 18 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem ciągłym – najlepsze wyniki dla zbioru o parametrach parametrów szumu 2x5%. ....	98
Tabela 19 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem ciągłym – najlepsze wyniki dla zbioru o parametrach parametrów szumu 2x10%. ....	98
Tabela 20 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem ciągłym – najlepsze wyniki dla zbioru o parametrach parametrów szumu 5x5%. ....	99
Tabela 21 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem ciągłym – najlepsze wyniki dla zbioru o parametrach parametrów szumu 10x5%. ....	99

Tabela 22 Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem przedziałowym (3 klasy; granica klasy – 15 tygodni) – najlepsze uzyskane wyniki. ....	101
Tabela 23 Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem przedziałowym (3 klasy; granica klasy – 35 tygodni) – najlepsze uzyskane wyniki. ....	101
Tabela 24 Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem przedziałowym (4 klasy; granica klasy – 15 tygodni) – najlepsze uzyskane wyniki. ....	102
Tabela 25 Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem przedziałowym (4 klasy; granica klasy – 35 tygodni) – najlepsze uzyskane wyniki. ....	103
Tabela 26 Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem binarnym (próg 35 tygodni) – najlepsze uzyskane wyniki. ....	104
Tabela 27 Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem binarnym (próg 35 tygodni) oraz funkcją Hampela jako funkcją kosztu – najlepsze uzyskane wyniki. ....	105
Tabela 28 Architektury testowane w trakcie doświadczeń na danych pierwotnych z wyjściem binarnym (próg 30 tygodni) – najlepsze uzyskane wyniki. ....	107
Tabela 29 Parametry zaszumiania danych; ilość rekordów zbiorów wynikowych – wyjście binarne, 30 wejść, próg 35 tygodni. ....	110
Tabela 30 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progu 35 tyg. – najlepsze wyniki dla zbioru o parametrach parametrów szumu 2x2%. ....	110
Tabela 31 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progu 35 tyg. – najlepsze wyniki dla zbioru o parametrach parametrów szumu 2x5%. ....	112
Tabela 32 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progu 35 tyg. – najlepsze wyniki dla zbioru o parametrach szumu 2x10%. ....	113
Tabela 33 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progu 35 tyg. – najlepsze wyniki dla zbioru o parametrach szumu 5x5%. ....	115
Tabela 34 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progu 35 tyg. – najlepsze wyniki dla zbioru o parametrach parametrów szumu 10x5%. ....	116
Tabela 35 Architektury testowane w trakcie doświadczeń na danych zaszumianych z wyjściem binarnym dla progu 30 tyg. – najlepsze wyniki dla zbioru o parametrach parametrów szumu 5x5%. ....	118
Tabela 36 Architektury testowane w trakcie doświadczeń na danych z wyjściem binarnym, opisanych przez 31 wejść dla domyślnej wartości progu - najlepsze wyniki. ....	122
Tabela 37 Architektury testowane w trakcie doświadczeń na danych z wyjściem binarnym, opisanych przez 31 wejść dla modyfikowanej wartości progu - najlepsze wyniki. ....	124
Tabela 38 Epidemiologiczne zbiory danych wykorzystane w pracy .....	127



Tabela 39 Zmienne opisujące rekordy pacjentek włączonych do badania (cukrzyca ciężarnych, 8 wejść).....	132
Tabela 40 Najlepsze uzyskane wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych pierwotnych – dane balansowane ujemnie.....	133
Tabela 41 Najlepsze uzyskane wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych pierwotnych – dane balansowane dodatnio.....	133
Tabela 42 Zakres zmian zmiennej <i>wiek</i> w problemie modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych przeformatowanych	134
Tabela 43 Zakres zmian zmiennej <i>rasa</i> w problemie modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych przeformatowanych	134
Tabela 44 Najlepsze uzyskane wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych przekodowanych – dane balansowane ujemnie.....	134
Tabela 45 Najlepsze uzyskane wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych przekodowanych – dane balansowane dodatnio.....	135
Tabela 46 Zmienne opisujące rekordy pacjentów włączonych do badania (nadciśnienie tętnicze, 38 wejść – po analizie wrażliwości).....	137
Tabela 47 Zmienne opisujące rekordy pacjentów włączonych do badania (nadciśnienie tętnicze, 8 wejść – po analizie wrażliwości.....	139
Tabela 48 Zmienne opisujące rekordy pacjentów włączonych do badania (nadciśnienie tętnicze, 6 wejść - ręczne) .....	140
Tabela 49 Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych pierwotnych (38 wejść) – dane balansowane ujemnie.....	141
Tabela 50 Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych pierwotnych – dane balansowane dodatnio.....	142
Tabela 51 Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych pierwotnych (8 wejść) – dane balansowane ujemnie.....	143
Tabela 52 Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych pierwotnych (8 wejść) – dane balansowane dodatnio.....	144
Tabela 53 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych pierwotnych (8 wejść), balansowanych dodatnio. ....	144
Tabela 54 Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych uzyskanych w oparciu o analizę literatury (6 wejść) – dane balansowane ujemnie. ....	145
Tabela 55 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych uzyskanych w oparciu o analizę literatury (6 wejść), balansowanych ujemnie.....	145
Tabela 56 Najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego z wykorzystaniem danych uzyskanych w oparciu o analizę literatury (6 wejść) – dane balansowane dodatnio. ....	146

Tabela 57 Wykresy ROC oraz wartości AUROC dla najlepszych modeli ryzyka nadciśnienia tętniczego uzyskane na danych uzyskanych w oparciu o analizę literatury (6 wejść), balansowanych dodatnio. ....	146
Tabela 58 Zmienne opisujące rekordy pacjentów włączonych do badania (atak serca, 28 wejść – po analizie wrażliwości) .....	149
Tabela 59 Zmienne opisujące rekordy pacjentów włączonych do badania (atak serca, 8 wejść - ręczne). .....	150
Tabela 60 Najlepsze uzyskane wyniki modelowania ryzyka ostrego incydentu wieńcowego z wykorzystaniem danych pierwotnych (28 wejść) – dane balansowane ujemnie. ....	151
Tabela 61 Najlepsze uzyskane wyniki modelowania ryzyka ostrego incydentu wieńcowego z wykorzystaniem danych pierwotnych (28 wejść) – dane balansowane dodatnio. ....	152
Tabela 62 Najlepsze uzyskane wyniki modelowania ryzyka ostrego incydentu wieńcowego z wykorzystaniem danych pierwotnych (8 wejść) – dane balansowane ujemnie. ....	153
Tabela 63 Najlepsze uzyskane wyniki modelowania ryzyka ostrego incydentu wieńcowego z wykorzystaniem danych pierwotnych (28 wejść) – dane balansowane dodatnio. ....	154
Tabela 64 Najlepsze uzyskane wyniki modelowania ryzyka wystąpienia choroby nowotworowej z wykorzystaniem danych pierwotnych (101 wejść) – dane balansowane ujemnie. ....	157
Tabela 65 Najlepsze uzyskane wyniki modelowania ryzyka wystąpienia choroby nowotworowej z wykorzystaniem danych pierwotnych (101 wejść) – dane balansowane dodatnio. ....	158
Tabela 66 Najlepsze uzyskane wyniki modelowania ryzyka wystąpienia choroby nowotworowej z wykorzystaniem danych pierwotnych (18 wejść) – dane balansowane ujemnie. ....	159
Tabela 67 Najlepsze uzyskane wyniki modelowania ryzyka wystąpienia choroby nowotworowej z wykorzystaniem danych pierwotnych (18 wejść) – dane balansowane dodatnio. ....	160
Tabela 68 Zamodelowana reakcja systemu (okres przeżycia pacjentów) na zmienioną (zredukowaną) dawkę chemioterapeutyku. ....	164
Tabela 69 Porównanie skuteczności chemioterapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (gemcytabina) oraz zamodelowanych (winorelbina). ....	166
Tabela 70 Porównanie skuteczności chemioterapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (winorelbina) oraz zamodelowanych (gemcytabina). ....	166
Tabela 71 Porównanie kosztów całkowitych dla schematów cis-platyna/winorelbina oraz cis-platyna/gemcytabina [120]. ....	167
Tabela 72 Średnie ceny porównywanych cytostatyków we wszystkich ośrodkach. ....	167
Tabela 73 Porównanie skuteczności chemioterapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (karboplatyna) oraz zamodelowanych (cis-platyna). ....	168

Tabela 74 Porównanie skuteczności chemioterapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (cisplatyna) oraz zamodelowanych (karboplatyna).....	168
Tabela 75 Porównanie wpływu zmiany leku antiemetycznego na wynik terapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (ondansetron) oraz zamodelowanych (metoklopramid). .....	170
Tabela 76 Porównanie wpływu zmiany leku antiemetycznego na wynik terapii (odsetek pacjentów przeżywających powyżej 35 tyg.) dla danych pierwotnych (metoklopramid) oraz zamodelowanych (ondansetron).....	170
Tabela 77 Porównanie najlepszego modelu neuronowego z modelami regresyjnymi - dane pierwotne z wyjściem binarnym (próg 35 tygodni) .....	175
Tabela 78 Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych pierwotnych z wyjściem binarnym (próg 35 tygodni) – różne metody modelowania.....	175
Tabela 79 Porównanie najlepszego modelu neuronowego z modelami regresyjnymi - dane zaszumiane (parametry szumu – 2x2%) z wyjściem binarnym (próg 35 tygodni). .....	177
Tabela 80 Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych zaszumianych (2x2%) z wyjściem binarnym (próg 35 tygodni) – różne metody modelowania.....	177
Tabela 81 Porównanie najlepszego modelu neuronowego z modelami regresyjnymi - dane zaszumiane (parametry szumu – 5x5%) z wyjściem binarnym (próg 35 tygodni). .....	179
Tabela 82 Wykresy ROC oraz wartości AUROC dla modeli uzyskanych w trakcie doświadczeń na danych zaszumianych (5x5%) z wyjściem binarnym (próg 35 tygodni) – różne metody modelowania.....	179
Tabela 83 Porównanie - najlepsze uzyskane wyniki modelowania ryzyka cukrzycy ciężarnych z wykorzystaniem danych przekodowanych – różne metody modelowania. ....	182
Tabela 84 Porównanie - najlepsze uzyskane wyniki modelowania ryzyka nadciśnienia tętniczego – różne metody modelowania. ....	185
Tabela 85 Porównanie - najlepsze uzyskane wyniki modelowania ryzyka ostrego incydentu wieńcowego – różne metody modelowania.....	187
Tabela 86 Porównanie - najlepsze uzyskane wyniki modelowania ryzyka wystąpienia choroby nowotworowej – różne metody modelowania. ....	189

## 15 Załączniki

### Załącznik 1.

Opis zmiennych sterujących zawartych w pierwotnym zbiorze BRFSS, 1999.

L.p.	Kod	Opis
1	_STATE	STATE FIPS CODE
2	_GEOSTR	STRATUM CODE
3	_DENSTR	DENSITY STRATUM
4	_PSU	PRIMARY SAMPLING UNIT
5	_RECORD	RECORD NUMBER
6	IMONTH (Character)	DATE OF INTERVIEW
7	IDAY (Character)	MONTH OF INTERVIEW
8	IYEAR (Character)	DAY OF INTERVIEW
9	INTVID (Character)	YEAR OF INTERVIEW
10	AREACODE (Character)	INTERVIEWER ID
11	PREFIX (Character)	PHONE NUMBER
12	SUFFIX (Character)	AREA CODE
13	DISPCODE	PREFIX
14	WINDDOWN	SUFFIX
15	SEQNO	DISPOSITION OF INTERVIEW
16	NATTMPTS	INTERVIEW IN REGULAR OR WINDDOWN MODE
17	NRECSEL	ANNUAL SEQUENCE NUMBER
18	NRECSTR	NUMBER OF ATTEMPTS
19	BPSELWT	NUMBER OF SAMPLE RECORDS FOR STRATUM
20	WAKSS1S2	NO. OF PH. NOS. IN STRATUM FROM SAMPLE
21	NUMADULT	BASIC PROBABILITY OF SELECTION WEIGHT
22	NUMMEN	WAKSBERG STAGE ONE OR STAGE TWO NUMBER
23	NUMWOMEN	NUMBER OF ADULTS IN HOUSEHOLD
24	GENHLTH	NUMBER OF ADULT MALES IN HOUSEHOLD
25	PHYSHLTH	NUMBER OF ADULT FEMALES IN HOUSEHOLD
26	MENTHLTH	GENERAL HEALTH
27	POORHLTH	NUMBER OF DAYS PHYSICAL HEALTH NOT GOOD
28	HLTHPLAN	NUMBER OF DAYS MENTAL HEALTH NOT GOOD
29	MEDICAR2	POOR PHYSICAL HEALTH
30	TYPCOVR1	HAVE ANY KIND OF HEALTH PLAN
31	TYPCOVR2	HAVE MEDICARE
32	NOCOV12	TYPE COVERAGE PAYING FOR MOST MED CARE
33	PASTPLAN	TYPE COVERAGE MAYBE NOT CONSIDERED
34	MEDCOST	NO HEALTH CARE COVERAGE IN PAST 12 MNTHS
35	CHECKUP	LAST HAD HEALTH COVERAGE
36	BPTAKE	TIME WHEN COULD NOT AFFORD TO SEE DOCTOR
37	BPHIGH	HOW LONG SINCE LAST ROUTINE CHECKUP

38	HIGHT1	HOW LONG SINCE BLOOD PRESSURE TAKEN
39	BLOODCHO	EVER TOLD BLOOD PRESSURE HIGH
40	CHOLCHK	TOLD MORE THAN ONCE BP HIGH
41	TOLDHI	EVER HAD BLOOD CHOLESTEROL CHECKED
42	DIABETES	HOW LONG SINCE CHOLESTEROL CHECKED
43	LASTDEN2	EVER TOLD BLOOD CHOLESTEROL HIGH
44	RMVTEETH	EVER TOLD BY DOCTOR YOU HAVE DIABETES
45	DENCLEAN	HOW LONG SINCE LAST DENTIST VISIT
46	SUNBURN	NUMBER OF PERMANENT TEETH REMOVED
47	NUMBURN	HOW LONG SINCE LAST TEETH CLEANING
48	SMOKE100	HAD A SUNBURN IN LAST 12 MONTHS
49	SMOKEDAY	NUMBER OF SUNBURNS
50	SMOKENUM	SMOKED AT LEAST 100 CIGARETTES
51	SMOKNM30	FREQUENCY OF DAYS NOW SMOKING
52	STOPSMOK	NUMBER OF CIGARETTES NOW SMOKED PER DAY
53	LASTSMOK	CIGS SMOKED WHEN SMOKED LT 30 DAYS
54	DRINKANY	QUIT SMOKING A DAY OR MORE IN PAST YEAR
55	ALCOHOL	HOW LONG SINCE LAST SMOKED CIGS
56	NALCOCC	ALCOHOLIC BEVERAGES IN PAST MONTH
57	DRINKGE5	HOW OFTEN DRINK LIQUOR
58	DRINKDRI	AVERAGE NUMBER OF DRINKS PER OCCASION
59	AGE	FIVE OR MORE DRINKS ON AN OCCASION
60	ORACE	DRINK AND DRIVE
61	HISPANIC	REPORTED AGE IN YEARS
62	MARITAL	ORIGINAL REPORTED RACE
63	CHLD04	HISPANIC ORIGIN
64	CHLD0512	MARITAL STATUS
65	CHLD1317	NUMBER OF CHILDREN LT 5 YEARS OLD
66	EDUCA	NUMBER OF CHILDREN 5 TO 12 YEARS OLD
67	EMPLOY	NUMBER OF CHILDREN 13 TO 17 YEARS OLD
68	INCOME2	EDUCATION LEVEL
69	WEIGHT	EMPLOYMENT STATUS
70	HEIGHT	INCOME LEVEL
71	CTYCODE	REPORTED WEIGHT IN POUNDS
72	NUMHHOLD	REPORTED HEIGHT IN FEET AND INCHES
73	NUMPHONS	COUNTY CODE
74	SEX	MORE THAN ONE TELEPHONE NUMBER IN HH
75	HADMAM	NUMBER OF UNIQUE PHONE NUMBERS IN HH
76	HOWLONG	SEX
77	WHYDONE	HAVE YOU EVER HAD A MAMMOGRAM
78	PROFEXAM	HOW LONG SINCE LAST MAMMOGRAM
79	LENGEXAM	REASON YOU HAD MAMMOGRAM
80	REASEXAM	EVER HAD BREAST PHYSICAL EXAM BY DOCTOR
81	HADPAP	HOW LONG SINCE LAST BREAST PHYSICAL EXAM
82	LASTPAP	REASON YOU HAD BREAST PHYSICAL EXAM
83	WHYPAP	EVER HAD A PAP SMEAR TEST

84	HADHYST	WHEN WAS LAST PAP SMEAR
85	PREGNANT	REASON YOU HAD PAP SMEAR
86	FLUSHOT	EVER HAD A HYSTERECTOMY
87	FLUPROV	ARE YOU NOW PREGNANT
88	PNEUMVAC	HAD FLU SHOT IN PAST 12 MONTHS
89	BLDSTOOL	KIND OF PLACE HAD FLU SHOT
90	LSTBLDST	EVER HAD PNEUMONIA VACCINATION
91	HADSIGM	EVER HAD BLOOD STOOL TEST USING HOME KIT
92	LASTSIGM	HOW LONG SINCE LAST HOME BLOOD STOOL TST
93	CHLDLT16	EVER HAD A SIGMOIDOSCOPY/COLONOSCOPY
94	BIKEHMT	HOW LONG SINCE SIGMOIDOSCOPY/COLONOSCOPY
95	SMKDETE2	AGE OF OLDEST CHILD UNDER AGE 16
96	GRADE2	BICYCLE HELMET WORN WHEN RIDING
97	CONDUSE	LAST TIME TESTED SMOKE DETECTORS
98	GETHIV	GRADE TO BEGIN HIV INFECTN AND AIDS EDUC
99	DONBLD85	ENCOURAGE SEX ACTIVE TEEN TO USE CONDOM
100	DON12MO	CHANCES OF GETTING INFECTED WITH HIV
101	HIVTST2A	DONATED BLOOD SINCE MARCH 1985
102	HIVTST2B	DONATED BLOOD IN PAST 12 MONTHS
103	HIV12MO	EVER TSTD FOR HIV OTHR THAN BLOOD DONATN
104	TEST12MO	EVER TESTED FOR HIV
105	RSNTST2	HIV TST PAST 12MO OTHR THAN BLOOD DONATN
106	WHRTST3	TESTED FOR HIV PAST 12 MONTHS
107	RESULTS	MAIN REASON FOR HIV TEST
108	COUNSEL	WHERE WAS LAST BLOOD TEST FOR HIV
109	DIABAGE	RECEIVED RESULTS OF LAST TEST
110	INSULIN	RECEIVED COUNSELING ABOUT TEST RESULTS
111	INSLNFRQ	AGE WHEN TOLD DIABETIC
112	BLDSUGAR	NOW TAKING INSULIN
113	HEMOGLBN	HOW OFTEN DO YOU USE INSULIN
114	DOCTDIAB	HOW OFTEN CHECK BLOOD FOR GLUCOSE
115	CHKHEMO	HEARD OF GLYCOSYLATED HEMOGLOBIN
116	FEETCHK	TIMES SEEN HEALTH PROF FOR DIABETES
117	EYEEXAM	TIMES CHECK FOR GLYCOSYLATED HEMOGLOBIN
118	VISNDIST	TIMES FEET CHECK FOR SORES/IRRITATIONS
119	VISNREAD	LAST EYE EXAM WHERE PUPILS WERE DILATED
120	VISIONTV	LIMITED VISION SEEING ACROSS STREET
121	SEXINTMN	LIMITED VISION IN READING PRINT
122	SEXCONDM	LIMITED VISION IN WATCHING TELEVISION
123	CONDLAST	HOW MANY PARTNERS IN PAST 12 MONTHS
124	CONEFF2	CONDOM USED LAST INTERCOURSE
125	NEWPARTN	WHY WAS CONDOM USED
126	HIVRISK	CONDOM EFFECTIVENESS AGAINST HIV
127	STDTRTAT	NUMBER NEW SEX PRTRNS IN LAST 12 MONTHS
128	STDCLIN	ANY HIGH RISK SITUATIONS APPLY
129	SEXBEHA2	TREATED FOR STD OR VD IN PAST FIVE YEARS

130	SELCPTN3	TREATED AT HEALTH DEPT STD CLINIC?
131	SEX1PTN3	CHANGED SEXUAL BEHAVIOR LAST 12 MOS.
132	USECOND3	DECREASE NUMB SEX PTNRS OR ABSTINENT
133	PREGLST5	NOW HAVE SEX WITH ONLY THE SAME PARTNER
134	LSTPREG	NOW ALWAYS USE CONDOMS FOR PROTECTION
135	LSTPREG2	PREGNANT IN LAST 5 YEARS
136	BRTHCNTL	FEEL ABOUT BECOMING PREGNANT-LAST
137	TYPCNTRL	FEEL ABOUT BECOMING PREGNANT-CURRENT
138	RSNNOBC	CURRENT BIRTH CONTROL USE
139	WHRSVCS	BIRTH CONTROL TYPE
140	FAMPLAN	REASON NO BIRTH CONTROL
141	LASTUSED	USUAL SOURCE FOR FEMALE HEALTH CARE
142	RSNNOCOV	EVER USE FAMILY PLANNING CLINIC
143	MCRELNGT	WHEN LAST USED CLINIC
144	DOCTLIST	MAIN REASON WITHOUT HEALTH CARE COVERAGE
145	CERTDOCT	HOW LONG HAD THIS HEALTH COVERAGE
146	HLTHCOVR	BOOK OR LIST OF DOCTORS WITH PLAN
147	RSNWOCOV	PLAN REQUIRES CERTAIN DOCTOR/CLINIC
148	RATECARE	ANY OTHER TYPE HEALTH CARE COVERAGE
149	PRIMCARE	REASON NO COVERAGE IN PAST TWELVE MONTHS
150	NOPRIMCR	RATE SATISFACTION WITH OVERALL CARE
151	MOSTCARE	PARTICULAR DOCTOR OR CLINIC
152	FACILIT2	MAIN REASON NO USUAL SOURCE OF CARE
153	DISTCARE	ONE PLACE YOU GO MOST OFTEN
154	NUMBDOC2	KIND OF PLACE YOU GO TO MOST OFTEN
155	CHNGDOCT	CONVENIENCE OF MEDICAL FACILITY LOCATION
156	LASTCHNG	ONE PARTICULAR DOCT FOR ROUTINE MED CARE
157	ASTHMA	LAST CHANGED DOCTORS
158	ASTHNOW	REASON CHANGED DOCTORS LAST TIME
159	PCSEATHB	EVER HAD ASTHMA
160	PCSEXER	HAVE ASTHMA NOW
161	PCSINJPV	DOCTOR TALKED ABOUT EATING HABITS
162	PCSDRGAB	DOC TALKED ABOUT EXERCISE OR PHYS ACTVTY
163	PCSALCH	DOCTOR TALKED ABOUT INJURY PREVENTION
164	PCSSMOK	DOCTOR TALKED ABOUT DRUG ABUSE
165	PCSAIDS	DOCTOR TALKED ABOUT ALCOHOL USE
166	CVDFATRK	DOCTOR TALKED ABOUT QUITTING SMOKING
167	CVDEXRSK	DOCTOR TALKED ABOUT SEXUAL PRACTICES
168	CVDFAT02	DOCTOR ADVISED FEWER FAT/CHOLESTEROL
169	CVDEXR02	DOCTOR ADVISED EXERCISE MORE
170	CVDINFAR	ARE YOU EATING FEWER FAT/CHOLESTEROL
171	CVDCORHD	ARE YOU EXERCISING MORE
172	CVDSTROK	EVER TOLD HAD MYOCARDIAL INFARCTION
173	CVDASPRN	EVER TOLD HAD CORONARY HEART DISEASE
174	ASPUNSAF	EVER TOLD HAD STROKE
175	WHYASPAN	TAKE ASPIRIN DAILY OR EVERY OTHER DAY

176	WHYASPHA	HEALTH MAKES TAKING ASPIRIN UNSAFE
177	WHYASPSK	TAKE ASPIRIN TO RELIEVE PAIN
178	PASTMENO	TAKE ASPIRIN TO REDUCE HEART ATTACK
179	DICSESTR	TAKE ASPIRIN TO REDUCE STROKE
180	ESTRPILL	GONE THROUGH MENOPAUSE
181	CURESTRO	DOC DISCUSSED BENEFITS/RISKS OF ESTROGEN
182	ESTRHART	DOCTOR PRESCRIBED ESTROGEN PILLS
183	ESTRBONE	CURRENTLY TAKING ESTROGEN
184	ESTRFLSH	TAKE/TOOK ESTROGEN TO PREVENT HRT ATTACK
185	PAIN12MN	TAKE/TOOK ESTROGEN FOR BONES
186	SYMTMMTH	TAKE/TOOK ESTROGEN TO TREAT MENOPAUSE
187	LMTJOINT	HAD PAIN, ACHING, STIFFNESS, SWELLING
188	HAVARTH	SYMPTOMS PRESENT AT LEAST ONE MONTH
189	TYPEARTH	LIMITED BECAUSE OF JOINT SYMPTOMS
190	TRTARTH	TOLD HAVE ARTHRITIS BY DOCTOR
191	FRUITJUI	TYPE ARTHRITIS DOCTOR DIAGNOSED
192	FRUIT	CURRENTLY TREATED BY DOC FOR ARTHRITIS
193	GREENSAL	HOW OFTEN DO YOU DRINK FRUIT JUICES
194	POTATOES	HOW OFTEN DO YOU EAT FRUIT
195	CARROTS	HOW OFTEN DO YOU EAT GREEN SALAD
196	VEGETABL	HOW OFTEN DO YOU EAT POTATOES
197	EXERANY	HOW OFTEN DO YOU EAT CARROTS
198	EXERACT1	SERVINGS OF VEGETABLES
199	EXERDIS1	PARTICIPATE IN ANY PHYSICAL ACTIVITY
200	EXEROFT1	TYPE OF PHYSICAL ACTIVITY
201	EXERHMM1	DISTANCE IN MILES AND TENTHS
202	EXEROTH2	HOW MANY TIMES ACTIVITY DONE
203	EXERACT2	HOW LONG ACTIVITY DONE
204	EXERDIS2	PARTICIPATE IN ANOTHER ACTIVITY
205	EXEROFT2	TYPE OF OTHER PHYSICAL ACTIVITY
206	EXERHMM2	DISTANCE IN MILES AND TENTHS
207	LOSEWT	HOW MANY TIMES SECOND ACTIVITY DONE
208	MAINTAIN	HOW LONG SECOND ACTIVITY DONE
209	FEWCAL	TRYING TO LOSE WEIGHT
210	PHYACT	TRYING TO MAINTAIN CURRENT WEIGHT
211	WTDESIRE	FEWER CALORIES AND/OR LESS FAT
212	DRADVICE	INCREASED PHYSICAL ACTIVITY
213	VITAMINS	DESIRED WEIGHT
214	MULTIVIT	PROFESSIONAL ADVICE ABOUT WEIGHT
215	FOLICACD	CURRENTLY TAKE VIT PILLS OR SUPPLEMENTS
216	TAKEVIT	ANY OF VITS/SUPPLEMENTS A MULTIVITAMIN
217	RECOMMEN	ANY VITAMINS CONTAIN FOLIC ACID
218	SUNBLOCK	HOW OFTEN TAKE VITAMIN OR SUPPLEMENT
219	SPFNUM	REASONS HLTH EXPERTS RECMND FOLIC ACID
220	AVOIDSUN	HOW OFTEN USE SUNSCREEN
221	SUNHAT	SPF OF SUNSCREEN



222	SUNWEAR	HOW OFTEN STAY IN SHADE
223	SUNSENS	HOW OFTEN WEAR HAT
224	BURNTYPE	HOW OFTEN WEAR LONG-SLEEVED SHIRTS
225	HOWSAFE	HOW SENSITIVE TO SUN
226	RENTHOME	HOW EASY TO SUNBURN
227	CURADRS	HOW SAFE IS NEIGHBORHOOD FROM CRIME
228	NMCLSFRD	RENT OR OWN HOME
229	ENGHFOOD	HOW LONG AT CURRENT ADDRESS
230	TOBACCO	NUMBR OF FRIENDS/CLOSE RELATIVES TO HELP
231	INDOORS	CNCRND ABOUT ENOUGH FOOD IN PAST 30 DAYS
232	SMKPUBLIC	SMOKING IN HOME PAST 30 DAYS
233	SMKWORK	INDOOR WORK LOCATION
234	SMKREST	SMOKING POLICY AT WORK FOR COMMON AREAS
235	SMKSCHLS	SMOKING POLICY AT WORK FOR COMMON AREAS
236	SMKDAYCR	SMOKING ALLOWED IN RESTAURANTS
237	SMKINDOR	SMOKING ALLOWED IN SCHOOLS
238	USEEVER	SMOKING ALLOWED IN DAY CARE CENTERS
239	USENOW	SMOKING ALLOWED IN INDOOR WORK AREAS
240	DISABLT	EVER USED SMOKELESS TOBACCO PRODUCTS
241	STATEQUE (Character)	CURRENTLY USE SMOKELESS TOBACCO PRODUCTS
242	Column 600	DISABILITY QUESTIONS
243	Column 601-750	STATE ADDED QUESTIONS
244	_QSTVER	QUESTIONNAIRE VERSION
245	_STSTR	SAMPLE DESIGN STRATIFICATION VARIABLE
246	_DENWT	DENSITY STRATUM WEIGHT
247	_GEOWT	GEOGRAPHIC STRATUM WEIGHT
248	_RAW	"RAW: # ADULTS IN HH/# PH #S TO HH"
249	_CSA	CLUSTER SIZE ADJUSTMENT
250	_WT1	DESIGN WEIGHT
251	_POSTSTR	POST STRATIFICATION WEIGHT
252	_FINALWT	FINAL WEIGHT
253	_REGION	REGION
254	_RACEG2_	RACE GROUP CODES USED IN POSTSTRTFICATION
255	_AGEG_	AGE GROUP CODES USED IN POSTSTRATFCATION
256	_SEXG_	GENDER GRP CODES USED IN POSTSTRTFICATION
257	_IMPAGE	IMPUTED AGE
258	_MSACODE	METROPOLITAN STATISTICAL AREA CODE
259	RACE	NEW RACE CODE
260	_SMOKER2	SMOKING STATUS(2)
261	_DRNKMO	TOTAL NUMBER DRINKS A MONTH
262	_WTFORHT	WEIGHT FOR HEIGHT PERCENT OF MEDIAN
263	_BMI	BODY MASS INDEX
264	_TOTINDX	PHYSICAL ACTIVITY LEVEL
265	_SMKLESS	SMOKELESS STATUS
266	_FRTSERV	DAILY SERVINGS OF FRUITS AND VEGETABLES
267	_FRTINDX	SUMMARY INDEX FOR FRUITS AND VEGETABLES

268	_RFHYPE2	HYPERTENSION(2)
269	_RFSMOK2	SMOKING(2)
270	_RFDRACU	ACUTE DRINKING
271	_RFDRDRI	DRINKING AND DRIVING
272	_RFDRCHR	CHRONIC DRINKING
273	_RFLIFES	SEDENTARY LIFESTYLE
274	_RFREGUL	REGULAR AND SUSTAINED PHYSICAL ACTIVITY
275	_RFOBESE	OBESITY (WEIGHT FOR HEIGHT % OF MEDIAN)
276	_RFWHBMI	OBESITY (BASED ON BMI)
277	_RFTOBAC	SMOKELESS TOBACCO (CURRENT USER)
278	_CHOLCHK	CHOLESTEROL CHECKED
279	_BPCHECK	BLOOD PRESSURE CHECKED
280	_AGEG	REPORTED AGE IN THREE AGE GROUPS
281	_AGEG5YR	REPORTED AGE IN 5 YEAR AGE CATEGORIES
282	_AGE65YR	REPORTED AGE IN TWO AGE GROUPS
283	_RACEG	NON-HISPANIC WHITES/ALL OTHERS
284	_RACEGR	RACE AND ETHNICITY

## Załącznik 2.

Dokładna charakterystyka danych wykorzystanych w trakcie modelowania ryzyka cukrzycy ciężarnych (wyselekcjonowane zmienne).

DIABETES	1	0	
	BPHIGH		suma
1	17	281	298
1 all	298		
%	<b>6%</b>	<b>94%</b>	
2	74	2179	2253
2 all	2253		
%	<b>3%</b>	<b>97%</b>	

DIABETES	1	0	
	TOLDHI		suma
1	8	215	223
1 all	223		
%	<b>4%</b>	<b>96%</b>	
2	39	1242	1281
2 all	1281		
%	<b>3%</b>	<b>97%</b>	

0	44	1003	1047
0 all	1047		
%	<b>4%</b>	<b>96%</b>	

DIABETES	<b>1</b>	<b>0</b>	
	SMOKEDAY		suma
1	12	563	575
1 all	575		
%	<b>2%</b>	<b>98%</b>	
2	7	111	118
2 all	118		
%	<b>6%</b>	<b>94%</b>	
3	14	307	321
3 all	321		
%	<b>4%</b>	<b>96%</b>	
0	58	1479	1537
0 all	1537		
%	<b>4%</b>	<b>96%</b>	

DIABETES	<b>1</b>	<b>0</b>	
	AGE		suma
mean	31.11	29.97	
SD	5.60	6.02	

DIABETES	<b>1</b>	<b>0</b>	
	ORACE		suma
1	76	1947	2023
1 all	2023		
%	<b>4%</b>	<b>96%</b>	
2	12	384	396
2 all	396		
%	<b>3%</b>	<b>97%</b>	
3	2	30	32
3 all	32		
%	<b>6%</b>	<b>94%</b>	
4	0	35	35
4 all	35		

%	<b>0%</b>	<b>100%</b>	
5	1	64	65
5 all	65		
%	<b>2%</b>	<b>98%</b>	

DIABETES	<b>1</b>	<b>0</b>	
	WEIGHT		suma
średnia	69.34	67.63	
SD	14.99	15.23	

DIABETES	<b>1</b>	<b>0</b>	
	HEIGHT		suma
średnia	163.70	165.04	
SD	7.69	6.86	

DIABETES	<b>1</b>	<b>0</b>	
	CVDFAT02		suma
1	30	837	867
1 all	867		
%	<b>3%</b>	<b>97%</b>	
2	17	451	468
2 all	468		
%	<b>4%</b>	<b>96%</b>	
0	44	1172	1216
0 all	1216		
%	<b>4%</b>	<b>96%</b>	

### Załącznik 3.

Opis poszczególnych zbiorów danych i zmiennych sterujących w kolejnych krokach modelowania ryzyka wystąpienia choroby nowotworowej (101 wejść, 18 wejść).

L.p.	Kod	Opis
1	HCFSR28	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CONDITION OF MENTAL ILLNESS (CODE 28)?

2	HCFSR19	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED DIGESTIVE/ENDOCRINE CONDITION (CODE 19)?
3	LVASECT	VASECTOMY REPORTED IN PREVIOUS INTERVIEW
4	HCFSR27	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED FEMALE REPRODUCTIVE CONDITION (CODE 27)?
5	LNAGE	SUBJECT'S AGE FROM PREVIOUS INTERVIEW
6	HCFSR24	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED MALE REPRODUCTIVE CONDITION (CODE 24)?
7	LCATRACT	CATARACTS REPORTED IN PREVIOUS INTERVIEW
8	LBKHIPYR	YEAR BROKEN HIP DIAGNOSED REPORTED IN PREVIOUS INTERVIEW
9	LOVARIES	BOTH OVARIES REPORTED REMOVED IN PREVIOUS INTERVIEW
10	MEDICAID	IS SUBJECT COVERED BY MEDICAID?
11	SURGRCP1	SUBJECT WAS RECUPERATING FROM SURGERY, ADMISSION #1
12	PREVMNHT	DURING THE MONTH BEFORE SUBJECT'S NURSING HOME ADMISSION OR DEATH WAS THERE A SIGNIFICANT CHANGE IN SUBJECT'S HEALTH?
13	HCFSR36	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CONDITION OF CATARACTS OTHER EYE DISORDERS (CODES 36 AND 3603)?
14	LLENOC25	LENGTH OF BIRTH CONTROL PILL USAGE BEFORE AGE 25 REPORTED IN
15	HCFSR15	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CONDITION OF SURGERY (CODE 15)?
16	QINTSPAN	WAS THE INTERVIEW CONDUCTED IN SPANISH FOR RESPONDENT
17	LNURSING	NURSING HOME RESIDENCE REPORTED IN PREVIOUS INTERVIEW
18	LOVARAGE	AGE WHEN LAST OVARY WAS REMOVED REPORTED IN PREVIOUS INTERVIEW
19	LSSEX	SUBJECT'S SEX
20	CORONBYP	HAS SUBJECT EVER HAD CORONARY BY-PASS SURGERY?
21	QBORING	DID QUESTIONNAIRE BORE RESPONDENT?
22	LCBPSURG	CORONARY BY-PASS SURGERY REPORTED IN

		PREVIOUS INTERVIEW
23	LDIABETE	DIABETES REPORTED IN PREVIOUS INTERVIEW
24	LPREG	AT LEAST ONE PREGNANCY REPORTED IN PREVIOUS INTERVIEW
25	VENTRAL	WAS THIS A VENTRAL HERNIA?
26	RUNDESST	DID RESPONDENT UNDERSTAND THE QUESTIONS?
27	HOSPFLU	WAS SUBJECT HOSPITALIZED FOR PNEUMONIA?
28	CNDTCNR1	CANCER REPORTED / ADMISSION #1
29	HOSPOTHR	HAS SUBJECT STAYED IN A HOSPITAL FOR ANY OTHER REASON?
30	SAVINGS	DID SUBJECT RECEIVE PERSONAL INCOME FROM INTEREST EARNINGS?
31	YEARSTOP	WHAT YEAR DID SUBJECT STOP SMOKING CIGARETTES?
32	SUBJCOND	DID DOCTOR SAY THAT THIS CONDITION WAS THE CAUSE OF SUBJECT'S DEATH?
33	QUPSTDEP	DID QUESTIONNAIRE UPSET OR DEPRESS RESPONDENT?
34	LPAPSMER	PAP SMEAR REPORTED IN PREVIOUS INTERVIEW
35	WHEELCHR	DOES SUBJECT USE A WHEELCHAIR?
36	LUTERAGE	AGE WHEN UTERUS WAS REMOVED REPORTED IN PREVIOUS INTERVIEW
37	HCFSR08	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CONDITION OF STROKE OR CVA (CEREBROVASCULAR ACCIDENT) (CODE 08)?
38	BRKNWRST	DID DOCTOR EVER SAY SUBJECT HAD A BROKEN WRIST?
39	EVERCATR	DID DOCTOR EVER SAY SUBJECT HAD CATARACTS?
40	CHLDSUP	DID SUBJECT RECEIVE PERSONAL INCOME FROM CHILD SUPPORT?
41	HERNIA	HAS A DOCTOR EVER SAY THAT SUBJECT HAD A HERNIA OR RUPTURE?
42	LWTAGE65	AGE 65 BODY WEIGHT REPORTED IN PREVIOUS INTERVIEW
43	HCFSR20	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED RESPIRATORY CONDITION (CODE 20)?
44	HCFSR26	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CIRCULATORY CONDITION (CODE 26)?
45	SUBDIENH	DID SUBJECT DIE IN HOSPITAL OR NURSING HOME?
46	MEDICARE	IS SUBJECT COVERED BY MEDICARE?
47	LKIDNEY	KIDNEY DISORDER/STONES REPORTED IN PREVIOUS

		INTERVIEW
48	AGEHPIL1	HOW OLD WAS SUBJECT WHEN SUBJECT FIRST TOOK HORMONE PILLS?
49	WOHRGAID	WITHOUT A HEARING AID, CAN SUBJECT HEAR AND UNDERSTAND WHAT A PERSON SAYS IN A NORMAL VOICE FROM ACROSS A QUIET ROOM?
50	SDIENRHM	WAS SUBJECT IN NURSING HOME OR OTHER HEALTH CARE FACILITY AT TIME OF DEATH?
51	LOSTEO	OSTEOPOROSIS REPORTED IN PREVIOUS INTERVIEW
52	HOSPNM1	ADDITIONAL HOSPITAL STAY REPORT - STAY #1
53	MANYPAP5	HOW MANY PAP TESTS DID SUBJECT HAVE IN LAST 5 YEARS?
54	LPHOSP	HOSPITALIZATIONS REPORTED IN PREVIOUS INTERVIEW
55	SPBATH	DOES SUBJECT USE SPECIAL DEVICES IN THE BATHROOM?
56	HCFSR01	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CONDITION OF ARTHRITIS (CODE 01)?
57	UNITCOLR	TIME PERIOD RESPONDED IN
58	PRIVATE	IS SUBJECT COVERED BY PRIVATE INSURANCE?
59	LFEMSTER	FEMALE STERILIZATION REPORTED IN PREVIOUS INTERVIEW
60	LWORKHST	RESPONDENT REPORTED WORKING IN ANY PREVIOUS INTERVIEW
61	STAYBED	DOES SUBJECT HAVE TO STAY IN BED MOST OF DAY?
62	LUTERUS	UTERUS REPORTED REMOVED IN PREVIOUS INTERVIEW
63	HOSPDIAB	WAS SUBJECT HOSPITALIZED FOR DIABETES?
64	DRINKYR	HAS SUBJECT HAD AT LEAST 1 DRINK OF BEER, WINE, OR LIQUOR IN PAST YEAR?
65	SWGHT65	WHAT WAS SUBJECT'S USUAL WEIGHT AT AGE 65?
66	LHEARAID	HEARING AID USE REPORTED IN PREVIOUS INTERVIEW
67	CFREQ	ACTIVITIES QUESTIONS DETERMINATOR
68	LBROKHIP	BROKEN/FRACTURED HIP REPORTED IN PREVIOUS INTERVIEW
69	LBRKWRST	BROKEN/FRACTURED WRIST REPORTED IN PREVIOUS INTERVIEW
70	OFTNWINE	DURING THE PAST YEAR HOW OFTEN DID SUBJECT DRINK WINE?
71	HOSPHBP	WAS SUBJECT HOSPITALIZED FOR HIGH BLOOD PRESSURE?
72	LSTROKE	STROKE REPORTED IN PREVIOUS INTERVIEW
73	HCFSR33	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY

		DATA TAPE WITH A REPORTED CONDITION NOT ELSEWHERE CODED (CODE 33)?
74	HEALTH	WHAT HAS BEEN HAPPENING TO HEALTH OF SUBJECT DURING THE PAST 12 MONTHS?
75	CONDYR1	YEAR OF ADMISSION STAY #1
76	LWRKDATE	DATE SUBJECT LAST REPORTED WORKING (YYMMDD) IN PREVIOUS
77	AGEHPILL	HOW OLD WAS SUBJECT WHEN SUBJECT LAST TOOK HORMONE PILLS?
78	JOINTS	WITHIN PAST YEAR, HAS SUBJECT HAD PAIN, SWELLING, STIFFNESS IN JOINTS ON MOST DAYS FOR AT LEAST ONE MONTH?
79	HEARNORM	CAN SUBJECT HEAR A PERSON TALKING IN A NORMAL VOICE FROM ACROSS THE ROOM?
80	HCHARFLG	ONE OR MORE FACILITY STAYS REPORTED: 1992 NHEFS
81	PATCHEVR	HAS SUBJECT EVER USED AN ESTROGEN OR HORMONE SKIN PATCH?
82	HOSPHOTH	WAS SUBJECT HOSPITALIZED FOR ANY OTHER HEART CONDITION?
83	HEAR Aid	HAS SUBJECT EVER WORN A HEARING AID?
84	SPCHAIR	DOES SUBJECT USE SPECIAL OR BUILT UP CHAIR OR TOILET?
85	HCFSR13	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CONDITION OF ANOTHER TYPE OF BONE FRACTURE BESIDES A HIP FRACTURE (CODE 13)?
86	HCFSR03	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CONDITION OF HEART ATTACK (CODE 03)?
87	HCFSR25	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CONDITION OF MUSCULOSKELETAL PROBLEM OR INJURY (CODE 25)?
88	EXERPROG	DOES SUBJECT EXERCISE OR PLAY SPORTS ON A REGULAR BASIS?
89	LARTGOUT	ARTHRITIS CAUSED BY GOUT REPORTED IN PREVIOUS INTERVIEW
90	SSRR	DID SUBJECT RECEIVE PERSONAL INCOME FROM SOCIAL SECURITY OR RAILROAD RETIREMENT?
91	DIVIDEND	DID SUBJECT RECEIVE PERSONAL INCOME FROM DIVIDENDS?
92	DOCKID	DID DOCTOR EVER SAY SUBJECT HAD A URINARY TRACT OR KIDNEY INFECTION MORE THAN THREE TIMES?



93	ATTKYR1	WHAT YEAR WAS THAT 1ST ADDITIONAL HEART ATTACK?
94	HCFSR18	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CONDITION OF TESTS/OBSERVATION/X-RAYS/PHYSICAL EXAM (CODE 18)?
95	CURRSX1	SEX OF CURRENT HOUSEHOLD MEMBER #1
96	LUTI	URINARY TRACT/KIDNEY INFECTION REPORTED IN PREVIOUS INTERVIEW
97	LONGPAP	HOW LONG HAS IT BEEN SINCE SUBJECT HAD PAP TEST (IN YEARS)?
98	LONEOVAR	ONE OVARY REPORTED REMOVED IN PREVIOUS INTERVIEW
99	LMOMDEAD	MOTHER REPORTED DECEASED IN PREVIOUS INTERVIEW
100	CALCIUM	HAS SUBJECT EVER REGULARLY TAKEN CALCIUM PILLS OR CALCIUM RICH ANTACIDS?
101	HCFSR09	DOES SUBJECT HAVE AT LEAST ONE STAY RECORD ON THE 1992 NHEFS HEALTH CARE FACILITY STAY DATA TAPE WITH A REPORTED CONDITION OF DIABETES (CODE 09)?

L.p.	Kod	Opis
1	LSSEX	SUBJECT'S SEX
2	LWORKHST	RESPONDENT REPORTED WORKING IN ANY PREVIOUS INTERVIEW
3	LARTGOUT	ARTHRITIS CAUSED BY GOUT REPORTED IN PREVIOUS INTERVIEW
4	LBROKHIP	BROKEN/FRACTURED HIP REPORTED IN PREVIOUS INTERVIEW
5	LKIDNEY	KIDNEY DISORDER/STONES REPORTED IN PREVIOUS INTERVIEW
6	LUTI	URINARY TRACT/KIDNEY INFECTION REPORTED IN PREVIOUS INTERVIEW
7	LLENOC25	LENGTH OF BIRTH CONTROL PILL USAGE BEFORE AGE 25 REPORTED IN
8	LNAGE	SUBJECT'S AGE FROM PREVIOUS INTERVIEW
9	BRKNWRST	DID DOCTOR EVER SAY SUBJECT HAD A BROKEN WRIST?
10	HIHERNIA	WAS THIS A HIATAL HERNIA OF THE DIAPHRAGM?
11	CATRSURG	HAS SUBJECT EVER HAD SURGERY FOR CATARACTS?
12	RESCNT	NUMBER OF NURSING HOME OR OTHER NON-HOSPITAL ADMISSIONS REPORTED
13	PREVYRHT	DURING THE 12 MONTHS PRIOR TO NURSING HOME ADMISSION OR DEATH WHAT WAS HAPPENING TO

		SUBJECT'S HEALTH?
14	OFTNBEER	DURING THE PAST YEAR HOW OFTEN DID SUBJECT DRINK BEER?
15	DIVIDEND	DID SUBJECT RECEIVE PERSONAL INCOME FROM DIVIDENDS?
16	OTHSUP	DID SUBJECT RECEIVE PERSONAL INCOME FROM SOME OTHER SOURCE?
17	BADSECT	WAS THERE A PARTICULARLY UPSETTING OR PROBLEMATIC SECTION FOR RESPONDENT?
18	PERSNM1	HOUSEHOLD MEMBER #1