

237/2007

**Raport Badawczy**

**RB/66/2007**

**Research Report**

**Próba zastosowania reguł  
decyzyjnych do predykcji  
szeregów czasowych**

**G. Szkatuła, M. Krawczak**

**Instytut Badań Systemowych  
Polska Akademia Nauk**

**Systems Research Institute  
Polish Academy of Sciences**



# **POLSKA AKADEMIA NAUK**

## **Instytut Badań Systemowych**

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 3810100

fax: (+48) (22) 3810105

Kierownik Pracowni zgłaszający pracę:  
Prof. dr hab. inż. J. Kacprzyk

Warszawa 2007

# Próba zastosowania reguł decyzyjnych do predykcji szeregów czasowych

## 1 Wprowadzenie

Rozwój systemów informatycznych ułatwił gromadzenie dużych zbiorów danych dotyczących różnych zjawisk, np. rynków finansowych. Ważne stało się nie tylko efektywne przechowywanie takich danych, ale także ich analiza, zdolność interpretacji i wyciągania potencjalnie użytecznych i zrozumiałych dla człowieka wniosków. Wykorzystanie pełnej informacji zawartej w tych zbiorach oraz sformułowanie ukrytej w nich wiedzy zwykle przekracza możliwości percepcji człowieka. W związku z tym powstało zapotrzebowanie na nowe metody i narzędzia informatyczne, które mogą pomóc człowiekowi w odkrywaniu wiedzy z danych. Duże znaczenie zyskały metody uczenia maszynowego wykorzystujące informacje zawarte w dostarczonym zbiorze przykładów.

Baza danych zawierająca historyczne dane może być wykorzystana do utworzenia tzw. *przykładów danych wejściowych* oraz odpowiadających im danych *wyjściowych*. Taka para  $\{\text{wejścia}_p, \text{wyjście}_p\}$  opisuje przykładowe zachowanie się określonego procesu. Zakłada się, że takie same lub podobne przyczyny (dane wejściowe) powodują taki sam lub podobny skutek opisany przez dane wyjściowe. Zasada ta daje nam podstawy do przypuszczenia, że pojawienie się takich samych lub podobnych danych opisujących wejścia spowoduje takie same lub podobne wyjście w przyszłości.

Zwykle zadania dotyczące predykcji szeregów czasowych należą do kategorii zadań aproksymacji funkcji. Szeregi czasowe są podstawą do scharakteryzowania obserwowanego złożonego zjawiska i do predykcji przyszłych stanów tego zjawiska. Poprzez obserwowane zjawisko rozumiemy tutaj złożony system, którego stan (lub stany) mogą być obserwowane w postaci szeregu czasowego (lub szeregów czasowych).

Predykcja oraz chęć (albo często potrzeba) zrozumienia przeszłości skłania nas do poszukiwania praw, które wyjaśniają istotę zachowania się obserwowanego zjawiska. W przypadku, gdy są znane deterministyczne równania opisujące zjawisko, to rozwiązanie tych równań pozwala określić stan zjawiska na podstawie warunków początkowych. W przypadku, gdy nie są znane równania opisujące obserwowane zjawisko, a jesteśmy zainteresowani predykcją stanów przyszłych, musimy określić zarówno reguły rządzące zjawiskiem (modelowanie matematyczne) jak i aktualny stan.

W tym raporcie ograniczymy się do przypadku drugiego, tzn. gdy nie są znane równania opisujące interesujące nas zjawisko, a poszukiwane reguły opisujące dynamikę tego zjawiska muszą być uzyskane z przebiegu historii tego zjawiska. Potraktujemy zadanie predykcji szeregów czasowych nie jako zadanie aproksymacji funkcji, lecz jako zadanie klasyfikacji danych.

W problemie klasyfikacji dysponujemy zbiorem obiektów lub przykładów - czyli zestawem danych wejściowych, należących do pewnych określonych klas, przy czym klasa obiektu jest określana jako wartość wyjściowa. Każdy przykład jest opisany przez skończoną liczbę cech (atrybutów) wraz z określonymi war-

tościami tych cech (wartościami atrybutów). Zbiory tak określonych przykładów z ich podziałem na klasy mogą stać się punktem wyjścia przy tworzeniu opisu rozpatrywanych klas (zjawisk) z zastosowaniem metod tzw. inteligencji komputerowej (lub tzw. sztucznej inteligencji). Metody sztucznej inteligencji, takie jak np. sztuczne sieci neuronowe pozwalają na konstrukcję modelu - typu black box - symulującego zjawisko; metody indukcyjnego uczenia maszynowego pozwalają na opis zjawiska w postaci skończonego zbioru reguł w postaci JEŻELI ... TO .... W obu przypadkach zależy nam na narzędziach pozwalających wyjaśnić obserwowane zjawisko, przy czym w przypadku reguł otrzymane przesłanki w sposób jawny opisują zasady podejmowania decyzji i zapis taki jest bardziej czytelny dla człowieka niż inne reprezentacje.

Wydaje się, że metody uczenia maszynowego na podstawie przykładów tworzące reguły decyzyjne mogą być zastosowane do predykcji szeregów czasowych, jednak po uprzedniej odpowiedniej wstępnej obróbce danych wejściowych. W tej pracy podjęto próbę przedstawienia zadania predykcji szeregu czasowego jako zadania uczenia maszynowego na podstawie przykładów i opisanie problemu predykcji w postaci zbioru reguł decyzyjnych. Opis zaproponowanego podejścia oraz sposób przygotowania danych (tzn. dobór cech, dyskretyzację wartości tych cech oraz ich zapis w postaci lingwistycznej) został omówiony dla danych testowych wygenerowanych w oparciu o funkcję Mackey'a-Glassa.

## 2 Opis zastosowanego podejścia

Załóżmy, że dysponujemy skończonym zbiorem przykładów  $U$  opisanych za pomocą zmiennych ilościowych, będących wynikiem pewnych pomiarów lub obserwacji. Chcąc zapisać przykłady zgodnie z metodologią zbiorów rozmytych, należy określić stopnie przynależności wartości cech wybranych do opisu tych przykładów do zbiorów rozmytych określonych przez wybrane wartości lingwistyczne.

Załóżmy, że mamy skończony, niepusty zbiór przykładów  $U = \{u_1, u_2, \dots, u_N\}$  i skończony, niepusty zbiór cech  $A = \{a_1, \dots, a_K\}$  opisujących te przykłady oraz  $V = \bigcup_{a \in A} V_a$  dla  $j = 1, \dots, K$ , gdzie  $V_a = \{v_{1(a)}^a, v_{2(a)}^a, \dots, v_{l(a)}^a\}$  jest niepustym, skończonym zbiorem wartości lingwistycznych cechy  $a_j \in A$ . Indeks  $l(j)$  określa liczbę wartości, jaką przyjmuje  $j$ -ta cecha. W procesie rozmywania dla każdej wartości cechy  $a_j$  określamy zbiór liczb  $\mu_{v_{1(a)}^a}(a_j), \dots, \mu_{v_{l(a)}^a}(a_j)$  z przedziału  $[0,1]$  określających, w jakim stopniu cecha  $a_j$  przyjmuje wartość lingwistyczną ze zbioru  $V_{a_j}$ .

Relacja rozmyta ze zbioru  $U$  do zbioru  $V_{a_j}$ , oznaczana  $\tilde{R}_{a_j}$ , określana jest na iloczynnie kartezjańskim  $U \times V_{a_j} \rightarrow [0,1]$  i definiowana w postaci zbioru uporządkowanego

$$\tilde{R}_{a_j} = \{(u, v); \mu_{\tilde{R}_{a_j}}(u, v) \mid \forall u \in U, \forall v \in V_{a_j}\} \quad (1)$$

gdzie  $\mu_{\tilde{R}_{a_j}}(u, v) \in [0,1]$  jest funkcją przynależności określającą, w jakim stopniu przykład  $u \in U$  jest w relacji z wartością  $v \in V_{a_j}$ . Relacja ta może być reprezentowana w postaci zbioru par albo w postaci macierzy relacji, której elementami są wartości funkcji przynależności określające, w jakim stopniu przykład  $u_i \in U$ ,  $i = 1, \dots, N$  jest w rozmytej relacji z wartością  $v^{a_j} \in V_{a_j}$ ,

	$v_1^{a_j}$	$v_2^{a_j}$	...	$v_{l(j)}^{a_j}$
$u_1$	$\mu_{\tilde{R}_{a_j}}(u_1, v_1^{a_j})$	$\mu_{\tilde{R}_{a_j}}(u_1, v_2^{a_j})$	...	$\mu_{\tilde{R}_{a_j}}(u_1, v_{l(j)}^{a_j})$
...	...	...	...	...
$u_N$	$\mu_{\tilde{R}_{a_j}}(u_N, v_1^{a_j})$	$\mu_{\tilde{R}_{a_j}}(u_N, v_2^{a_j})$	...	$\mu_{\tilde{R}_{a_j}}(u_N, v_{l(j)}^{a_j})$

lub też w postaci rozmytego skierowanego grafu, którego wierzchołki stanowią elementy rozważanych zbiorów, krawędzie odpowiadają elementom relacji, to znaczy uporządkowanym parom, a waga każdej skierowanej krawędzi odpowiada wartości odpowiedniej funkcji przynależności.

Niech  $U$  będzie skończonym, niepustym zbiorem przykładów,  $A$  niepustym zbiorem cech opisujących te przykłady. Niech cecha  $a_j \in A$  ma zbiór wartości  $V_{a_j} = \{v_1^{a_j}, v_2^{a_j}, \dots, v_{l(j)}^{a_j}\}$ . Zbiorem rozmytym w zbiorze  $U$  określonym przez wartość  $v_i^{a_j} \in V_{a_j}$ , oznaczanym  $\tilde{U}_{a_j=v_i^{a_j}}$ , nazywamy podzbiór relacji rozmytej  $\tilde{R}_{a_j}$  ze zbioru  $U$  do  $V_{a_j}$  dla ustalonej wartości  $v_i^{a_j} \in V_{a_j}$ , zapisywany w postaci

$$\tilde{U}_{a_j=v_i^{a_j}} = \{(u, v_i^{a_j}); \mu_{\tilde{R}_{a_j}}(u, v_i^{a_j}) \mid \forall u \in U, \mu_{\tilde{R}_{a_j}}(u, v_i^{a_j}) > 0\}.$$

Zbiór rozmyty określony poprzez warunek stanowi uporządkowany zbiór trójek:  $(u, v_i^{a_j})$  i wartości funkcji przynależności  $\mu_{\tilde{R}_{a_j}}(u, v_i^{a_j})$  określającej, w jakim stopniu w przykładzie  $u$  rozpatrywana cecha  $a_j$  przyjmuje wartość  $v_i^{a_j}$ .

Jeżeli ze zbioru cech  $A$  wybierzemy jedną cechę  $a_d$ , to możemy ze względu na wartości, które ona przyjmuje dokonać rozmytego podziału zbioru przykładów  $U$ . Elementy zbioru  $A \setminus \{a_d\}$  nazywamy *cechami warunkowymi*, a cechę  $a_d$  nazywamy *cechą decyzyjną*.

Podziałem zbioru przykładów  $U$  ze względu na cechę decyzyjną  $a_d \in A$  ze zbiorem wartości  $V_d = \{v_1^{a_d}, \dots, v_{l(d)}^{a_d}\}$  nazywamy rozmyte zbiory przykładów

$$\tilde{U}_{a_d} = \{\tilde{U}_{a_d=v_1^{a_d}}, \tilde{U}_{a_d=v_2^{a_d}}, \dots, \tilde{U}_{a_d=v_{l(d)}^{a_d}}\},$$

gdzie  $\tilde{U}_{a_d=v_i^{a_d}} = \{(u, v_i^{a_d}); \mu_{\tilde{R}_{a_d \rightarrow v_i^{a_d}}}(u, v_i^{a_d}) | u \in U, \mu_{\tilde{R}_{a_d \rightarrow v_i^{a_d}}}(u, v_i^{a_d}) > 0\}$ .

$\tilde{U}_{a_d=v_i^{a_d}}$  jest zbiorem rozmytym w zbiorze  $U$  określonym przez wartość  $v_i^{a_d} \in V_{a_d}$  będącym podzbiorem relacji rozmytej  $\tilde{R}_{a_d}$  ze zbioru  $U$  do  $V_{a_d}$  dla ustalonej wartości  $v_i^{a_d} \in V_{a_d}$ . Funkcja przynależności  $\mu_{\tilde{R}_{a_d \rightarrow v_i^{a_d}}}(u, v_i^{a_d})$  dla  $u \in U$  określa, w jakim stopniu w przykładzie  $u$  cecha  $a_d$  przyjmuje wartość  $v_i^{a_d}$ .

Dla uproszczenia zadania uczenia, przyjmuje się w literaturze dwa dodatkowe wymagania, które muszą być spełnione przez zbiór przykładów ze zbioru  $U$ , przedstawione poniżej,

$$t(\mu_{\tilde{R}_{a_d \rightarrow v_{i_1}^{a_d}}}(u, v_{i_1}^{a_d}), \mu_{\tilde{R}_{a_d \rightarrow v_{i_2}^{a_d}}}(u, v_{i_2}^{a_d})) = 0, \text{ dla } i_1 \neq i_2, \forall u \in U \quad (2)$$

$$s(\mu_{\tilde{R}_{a_d \rightarrow v_{i_1}^{a_d}}}(u, v_{i_1}^{a_d}), \dots, \mu_{\tilde{R}_{a_d \rightarrow v_{i_k}^{a_d}}}(u, v_{i_k}^{a_d})) = 1, \text{ dla } \forall u \in U. \quad (3)$$

Interpretacja powyższych warunków jest następująca, dla każdego przykładu  $u \in U$  i dla każdej cechy  $a_j \in A$  mającej zbiór wartości  $V_{a_j} = \{v_1^{a_j}, v_2^{a_j}, \dots, v_{l(j)}^{a_j}\}$  istnieje dokładnie jedna wartość  $v_i^{a_j} \in V_{a_j}$  dla której funkcja przynależności  $\mu_{\tilde{R}_{a_j}}(u, v_i^{a_j}) = 1$  oraz dla  $\forall v \neq v_i^{a_j}, \mu_{\tilde{R}_{a_j}}(u, v) = 0$ .

Jeżeli w zadaniach praktycznych zbiory przykładów nie spełniają warunków (2) i (3), dokonuje się modyfikacji zbioru cech  $A$  wybranych do opisu przykładów, wprowadzając nowe cechy nominalne.

Tak określony zbiór rozmyty  $\tilde{U}_{a_d=v_i^{a_d}}$  będzie oznaczał klasę wyznaczoną przez rozmytą partycję, której opisu szukamy w postaci reguł typu *JEŻELI ... TO ...*

Każdy przykład  $u_n \in U$ ,  $n = 1, \dots, N$  opisany za pomocą cech  $a_j \in A$ ,  $j = 1, \dots, K$  możemy zapisać w następujący sposób

$$u_n : (a_1 = v_{t(1,n)}^{a_1}); \mu_{a_1=v_{t(1,n)}^{a_1}} \wedge (a_2 = v_{t(2,n)}^{a_2}); \mu_{a_2=v_{t(2,n)}^{a_2}} \wedge \dots \wedge (a_K = v_{t(K,n)}^{a_K}); \mu_{a_K=v_{t(K,n)}^{a_K}} \quad (4)$$

gdzie indeks  $t(j,n)$  określa wartość jaką przyjmuje  $j$ -ta cecha w  $n$ -tym przykładzie. Funkcja  $\mu_{a_j=v_{t(j,n)}^{a_j}}$  określa, w jakim stopniu w przykładzie  $u_n$  cecha  $a_j$  przyjmuje wartość  $v_{t(j,n)}^{a_j}$ .

W dalszej części pracy dla uproszczenia zapisu będziemy pomijać górne indeksy przy wartościach cechy decyzyjnej i będziemy stosować uproszczony zapis  $V_{a_d} = \{v_1, \dots, v_n\}$  oraz  $\{\tilde{U}_{v_1}, \tilde{U}_{v_2}, \dots, \tilde{U}_{v_{|a_d|}}\}$ , zatem każdy przykład  $u_n \in U$  można zapisać w postaci:

$$u_n : (a_1 = v_{i(1,n)}); \mu_{v_{i(1,n)}} \wedge (a_2 = v_{i(2,n)}); \mu_{v_{i(2,n)}} \wedge \dots \wedge (a_K = v_{i(K,n)}); \mu_{v_{i(K,n)}} \quad (5)$$

Tak zapisany zbiór przykładów jest punktem wyjścia w procesie uczenia maszynowego na podstawie przykładów, w którym otrzymujemy zbiór reguł rozmytych typu JEŻELI ... TO .... Reguła rozmyta może zostać zapisana w postaci

$$k\text{-ta reguła: JEŻELI } (a_1 = v_{i(1,k)}) \wedge \dots \wedge (a_K = v_{i(K,k)}) \text{ TO } (a_d = v_{i(d,k)}) \quad (6)$$

indeks  $l(j,k)$  określa wartość jaką przyjmuje cecha  $a_j$  w  $k$ -tej regule,  $k=1, \dots, S$ .

Disponując zbiorem  $k$ -reguł rozmytych możemy stosując wnioskowanie rozmyte obliczyć stopień spełnienia  $k$ -tej reguły przez każdy przykład  $u_n$  ze zbioru  $U$ .

Niech funkcja  $\mu_{i(j,k,n)}$  oznacza funkcję przynależności  $j$ -tej cechy w  $n$ -tym przykładzie do zbioru rozmytego określonego przez warunek  $(a_j = v_{i(j,k)})$  z części przesłankowej  $k$ -tej reguły. Po pierwsze, dla każdego przykładu  $u_n \in U$ ,  $n=1, \dots, N$  i każdej reguły możemy obliczyć wartości funkcji przynależności do zbiorów rozmytych z części przesłankowej reguły:  $\mu_{i(1,k,n)}, \mu_{i(2,k,n)}, \dots, \mu_{i(K,k,n)}$  oraz do zbioru określonego przez konkluzję:  $\mu_{i(d,k,n)}$ . Następnie, aby otrzymać stopień spełnienia przez przykład  $u_n$  części przesłankowej  $k$ -tej reguły stosujemy operację agregacji zbiorów rozmytych z części przesłankowej reguły i otrzymujemy

$$\tau_k(u_n) = A'(\mu_{i(1,k,n)}, \mu_{i(2,k,n)}, \dots, \mu_{i(K,k,n)}), \quad (7)$$

gdzie  $A$  jest wybranym operatorem koniunkcyjnym agregacji, zazwyczaj minimum.

Niech  $\tilde{C}_k$  określa klasę rozmytą określoną przez  $k$ -tą regułę. Stopień spełnienia  $k$ -tej reguły przez przykład  $u_n$  wynosi

$$\mu_{\tilde{C}_k}(u_n) = \min\{\tau_k(u_n), \mu_{i(d,k,n)}\}. \quad (8)$$

Agregując dla danego przykładu  $u_n$  zbiory rozmyte z konkluzji reguł opisujących ten przykład otrzymujemy wartość funkcji przynależności przykładu do jednej klasy  $\tilde{C}$

$$\mu_{\tilde{C}}(u) = A'(\mu_{\tilde{C}_1}(u), \mu_{\tilde{C}_2}(u), \dots, \mu_{\tilde{C}_S}(u)), \quad (9)$$

gdzie  $A$  jest wybranym operatorem dysjunkcyjnym agregacji, zazwyczaj maksimum.

### 3 Opis szeregu czasowego

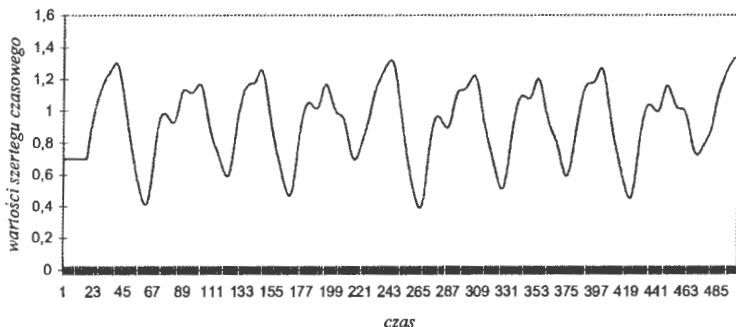
Analiza szeregów czasowych ma trzy zasadnicze cele: prognozowanie, modelowanie oraz analizę właściwości. Celem prognozowania (predykcji) jest próba dokładnego przewidywania zmiany stanu systemu, dla krótkiego horyzontu czasowego. Modelowanie ma na celu znalezienie opisu zmian stanu systemu w długim horyzontie czasowym. Trzeci cel analizy szeregów czasowych służy do określenia fundamentalnych właściwości, takich jak np. liczba stopni swobody systemu.

Za początek nowoczesnych metod prognozowania można przyjąć prace Yulea (1927), który zaproponował metodę *autoregresji* do prognozowania liczby plam na słońcu. W modelu Yulea prognoza przyszłej liczby plam była ważoną sumą poprzednio zaobserwowanych ilości plam na słońcu. Przez wiele lat liniowy model Yulea, uzupełniony przez szum, był podstawową metodą używaną do predykcji szeregów czasowych.

Rozważmy następujące równanie Mackey'a-Glassa (1977):

$$x(t+1) = bx(t) + a \frac{x(t-s)}{1+x^c(t-s)}. \quad (10)$$

Model systemu (10) jest całkowicie deterministyczny, tzn. rozwój systemu jest całkowicie określony przez warunki początkowe. Dla przykładowych parametrów:  $a=0.2$ ,  $b=0.9$ ,  $c=10$ ,  $s=18$  oraz warunków początkowych  $x(0)=x(1)=\dots=x(18)=0.7$  przebieg czasowy równania (10) przedstawiony został na Rysunku 1.



Rysunek 1. Pierwsze 500 punktów równania Mackey'a-Glassa

Poszczególne chwile  $k=1, 2, \dots, 500$  oraz odpowiadające im wartości szeregu czasowego  $x(k)$  tworzą party  $\{k, x(k)\}$ , jak pokazano w Tabeli 1 (dla  $\chi=19$ ):



Tabela 1.

$k$	$x(k)$
$\chi$	0,70000
$\chi+1$	0,76615
$\chi+2$	0,82569
$\chi+3$	0,87928
$\chi+4$	0,92750
$\chi+5$	0,97091
$\chi+6$	1,00997
$\chi+7$	1,04513
$\chi+8$	1,07677
$\chi+9$	1,10525
$\chi+10$	1,13088
$\chi+11$	1,15394
$\chi+12$	1,17470
$\chi+13$	1,19339
$\chi+14$	1,21020
$\chi+15$	1,22533
$\chi+16$	1,23896
$\chi+17$	1,25121
$\chi+18$	1,26225
$\chi+19$	1,27218
$\chi+20$	1,28821
$\chi+21$	1,30332
$\vdots$	$\vdots$

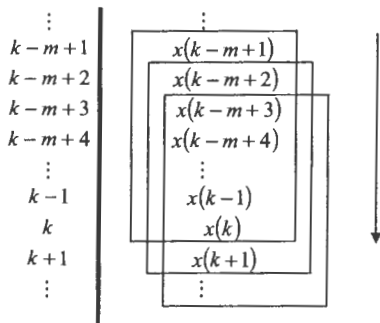
#### 4 Eksperyment obliczeniowy

Załóżmy, że dysponujemy skończonym zbiorem par  $\{k, x(k)\}$ , gdzie  $k$  oznacza chwile czasowe a  $x(k)$  odpowiadające im wartości szeregu czasowego (Tabela 1).

##### 4.1. Dane Nr 1

Załóżmy, że rozważamy okresy w szeregu czasowym o długości „okna” równej  $m$ , Tabela 2.

Tabela 2. „Okna o długości  $m$



Utwórzmy następujący zbiór przykładów (okien) przedstawiony w Tabeli 3:

Tabela 3

$k$	$x(k-m+1)$	$x(k-m+2)$	$x(k-m+3)$	...	$x(k-1)$	$x(k)$	$x(k+1)$
$\chi$	0,70000	0,76615	0,82569	...	1,19339	1,21020	<b>1,22533</b>
$\chi+1$	0,76615	0,82569	0,87928	...	1,21020	1,22533	<b>1,23896</b>
$\chi+2$	0,82569	0,87928	0,92750	...	1,22533	1,23896	<b>1,25121</b>
$\chi+3$	0,87928	0,92750	0,97091	...	1,23896	1,25121	<b>1,26225</b>
$\chi+4$	0,92750	0,97091	1,00997	...	1,25121	1,26225	<b>1,27218</b>
$\chi+5$	0,97091	1,00997	1,04513	...	1,26225	1,27218	<b>1,28821</b>
$\vdots$							

Każde  $m$  kolejnych wartości szeregu czasowego oznaczamy jako:

$$\mathbf{x}(k)=[x(k-m+1), x(k-m+2), x(k-m+3), \dots, x(k-1), x(k)]$$

zaś

$$y(k)=x(k+1)$$

dla  $k = \chi, \chi+1, \dots, \chi+M$ , gdzie  $M$  oznacza liczbę przykładów (okien) które uwzględniają  $M+m+1$  wartości szeregu czasowego (pomiarów).

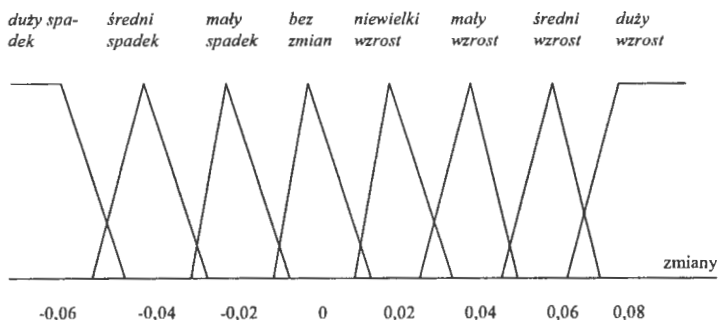
Tak utworzone pary  $\{\mathbf{x}(k), y(k)\}$ ,  $k = \chi, \chi+1, \dots, \chi+M$ , opisujące okna, mogą stanowić zbiór przykładów uczących dla zadania uczenia maszynowego.

Załóżmy, że dysponujemy skończonym zbiorem par (przykładów)  $\{\mathbf{x}(k), y(k)\}$ ,  $k = \chi, \chi+1, \dots, \chi+M$ , opisanych za pomocą zmiennych ilościowych, będących wynikiem pewnych pomiarów (Tabela 3).

Pierwszym krokiem jest zapisanie tych przykładów zgodnie z metodologią zbiorów rozmytych (tzw. rozmywanie), czyli określenie stopni przynależności wartości do odpowiednich zbiorów rozmytych, określonych przez wybrane wartości lingwistyczne:

- 1: „duży spadek”, 2: „średni spadek”, 3: „mały spadek”,  
 4: „bez zmian”,  
 5: „niewielki wzrost”, 6: „mały wzrost”, 7: „średni wzrost” i 8: „duży wzrost”.

Przyjęte funkcje przynależności do rozpatrywanych zbiorów rozmytych przedstawiono na Rys.2.



Rysunek 2. Funkcje przynależności

W Tabeli 4 przedstawiono praktyczny, uproszczony sposób określania przynależności w stopniu jeden, do wyżej wymienionych zbiorów rozmytych, dla wybranych 20 kolejnych wartości szeregu czasowego generowanego przez równanie (10). W kolumnie drugiej przedstawiono szereg generowany przez równanie (10)  $x(k)$ ; w kolumnie trzeciej różnice w chwilach  $x(k) - x(k-1)$ ; w następnej różnice wartości zostały pomnożone przez 100; zaś następnie w celu otrzymania dodatnich wartości szeregu dodano liczbę 8; w ostatniej kolumnie poprzednio otrzymane wartości podzielono przez 2 i zaokrąglono.

Tabela 4.

$k$	$x(k)$	$x(k) - x(k-1)$	$[x(k) - x(k-1)] * 100$	$[x(k) - x(k-1)] + 8$	$\{[x(k) - x(k-1)] + 8\} / 2$
1	0,70	0,00	0	8	4
2	0,77	0,07	7	15	7
3	0,83	0,06	6	14	7
4	0,88	0,05	5	13	7
5	0,93	0,05	5	13	6
6	0,97	0,04	4	12	6
7	1,01	0,04	4	12	6

8	1,05	0,04	4	12	6
9	1,08	0,03	3	11	6
10	1,11	0,03	3	11	5
11	1,13	0,03	3	11	5
12	1,15	0,02	2	10	5
13	1,17	0,02	2	10	5
14	1,19	0,02	2	10	5
15	1,21	0,02	2	10	5
16	1,23	0,02	2	10	5
17	1,24	0,01	1	9	5
18	1,25	0,01	1	9	5
19	1,26	0,01	1	9	5
20	1,27	0,01	1	9	4

Zakładamy, że stan  $y(k+1)$  (ozn.  $x(k+1)$ ) zależy od stanów poprzednich, tzn.

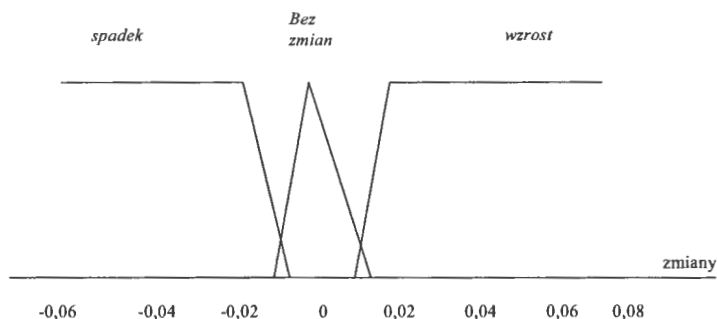
$$y(k+1) = \Psi[ (k)] \quad (11)$$

gdzie funkcja  $\Psi$  jest nieznaną funkcją stanów poprzednich, zaś  $m=15$  oznacza rozpatrywaną historię szeregu (długość rozpatrywanego okna).

Na tym etapie należy określić cechę decyzyjną (występującą we wnioskach reguł), której wartości będą określać klasy. Taką cechą decyzyjną może być cecha  $a_p$ : „zmiana w  $k+1$  kroku” z wartościami lingwistycznymi:

- 1: „spadek”,
- 2: „bez zmian”,
- 3: „wzrost”.

Przyjęte funkcje przynależności do rozpatrywanych zbiorów rozmytych przedstawiono na Rys.3.



Rysunek 3. Funkcje przynależności

W Tabeli 5 przedstawiono uproszczony, praktyczny sposób określania przynależności w stopniu jeden, do wyżej wymienionych trzech zbiorów rozmytych, dla wybranych 20 wartości szeregu czasowego generowanego przez równanie (10). W kolumnie drugiej przedstawiono szereg generowany przez równanie (10)  $x(k)$ ; w kolumnie trzeciej różnice w chwilach  $x(k) - x(k-1)$ ; w następnej różnice wartości zostały pomnożone przez 100; zaś następnie w celu otrzymania dodatnich wartości szeregu dodano liczbę 8 i podzielono na 2; w ostatniej kolumnie zamieszczono wartości lingwistyczne cechy decyzyjnej.

Tabela 5.

$k$	$x(k)$	$x(k) - x(k-1)$	$[x(k) - x(k-1)] * 100$	$\{[x(k) - x(k-1)] + 8\} / 2$	1 if $3,5 \leq (5)$ 2 if $3,5 < (5) < 4,5$ 3 if $(5) \geq 4,5$
(1)	(2)	(3)	(4)	(5)	(6)
1	0,70	0,00	0	4	2
2	0,77	0,07	7	7	3
3	0,83	0,06	6	7	3
4	0,88	0,05	5	7	3
5	0,93	0,05	5	6	3
6	0,97	0,04	4	6	3
7	1,01	0,04	4	6	3
8	1,05	0,04	4	6	3
9	1,08	0,03	3	6	3
10	1,11	0,03	3	5	3
11	1,13	0,03	3	5	3
12	1,15	0,02	2	5	3
13	1,17	0,02	2	5	3
14	1,19	0,02	2	5	3
15	1,21	0,02	2	5	3
16	1,23	0,02	2	5	3
17	1,24	0,01	1	5	3
18	1,25	0,01	1	5	3
19	1,26	0,01	1	5	3
20	1,27	0,01	1	4	3
21	1,29	0,02	2	5	3
22	1,30	0,02	2	5	3
23	1,31	0,01	1	4	2
24	1,31	0,00	0	4	2
25	1,29	-0,02	-2	3	1

Wartości lingwistyczne przedstawione w Tabeli 4 i Tabeli 5 zastosowano do przykładów zawartych w Tabeli 3, otrzymując zbiór przykładów uczących, zamieszczonych w Tabeli 6, zgodnie z przyjętymi oznaczeniami. Przy czym cecha  $a_1$  oznacza wartość szeregu  $x(k-14)$  w chwili  $(k-14)$ , cecha  $a_2$  wartość szeregu  $x(k-13)$  w chwili  $(k-13)$  itd., a cecha  $a_{15}$  oznacza wartość szeregu  $x(k)$  w chwili  $(k)$ . Cecha decyzyjna  $a_d$  oznacza wartość szeregu  $x(k+1)$  w chwili  $(k+1)$ .

Tabela 6. Zbiór przykładów uczących

Lp.	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	a15	$a_d$
u1	7	6	6	6	6	6	5	5	5	5	5	5	5	5	5	3
u2	6	6	6	6	6	5	5	5	5	5	5	5	5	5	5	2
u3	6	6	6	6	5	5	5	5	5	5	5	5	5	5	4	3
u4	6	6	6	5	5	5	5	5	5	5	5	5	5	4	5	3
u5	6	6	5	5	5	5	5	5	5	5	5	5	4	5	5	2
u6	6	5	5	5	5	5	5	5	5	5	5	4	5	5	4	2
u7	5	5	5	5	5	5	5	5	5	5	4	5	5	4	4	1
u8	5	5	5	5	5	5	5	5	5	4	5	5	4	4	3	1
u9	5	5	5	5	5	5	5	5	4	5	5	4	4	3	2	1
u10	5	5	5	5	5	5	5	4	5	5	4	4	3	2	2	1
u11	5	5	5	5	5	5	4	5	5	4	4	3	2	2	1	1
u12	5	5	5	5	5	4	5	5	4	4	3	2	2	1	1	1
u13	5	5	5	5	4	5	5	4	4	3	2	2	1	1	1	1
u14	5	5	5	4	5	5	4	4	3	2	2	1	1	1	1	1
u15	5	5	4	5	5	4	4	3	2	2	1	1	1	1	1	1
u16	5	4	5	5	4	4	3	2	2	1	1	1	1	1	1	1
u17	4	5	5	4	4	3	2	2	1	1	1	1	1	1	1	1
u18	5	5	4	4	3	2	2	1	1	1	1	1	1	1	1	1
u19	5	4	4	3	2	2	1	1	1	1	1	1	1	1	2	1
u20	4	4	3	2	2	1	1	1	1	1	1	1	1	2	2	1

Wartości cech zawarte w Tabeli 6 mają zgodnie z przyjętymi oznaczeniami interpretację lingwistyczną, jak pokazano w Tabeli 7.

Tabela 7. Zbiór przykładów uczących

Lp.	a1	...	a14	a15	$a_d$
u1	„średni wzrost”	...	„niewielki wzrost”,	„niewielki wzrost”,	„wzrost”
u2	„mały wzrost”,	...	„niewielki wzrost”,	„niewielki wzrost”,	„bez zmian”
u3	„mały wzrost”,	...	„niewielki wzrost”,	„bez zmian”,	„wzrost”
u4	„mały wzrost”,	...	„bez zmian”,	„niewielki wzrost”,	„wzrost”
u5	„mały wzrost”,	...	„niewielki wzrost”,	„niewielki wzrost”,	„bez zmian”
u6	„mały wzrost”,	...	„niewielki wzrost”,	„bez zmian”,	„bez zmian”
u7	„niewielki wzrost”,	...	„bez zmian”,	„bez zmian”,	„spadek”
u8	„niewielki wzrost”,	...	„bez zmian”,	„mały spadek”,	„spadek”
u9	„niewielki wzrost”,	...	„mały spadek”,	„średni spadek”,	„spadek”
u10	„niewielki wzrost”,	...	„średni spadek”,	„średni spadek”,	„spadek”
⋮					

Na podstawie zbioru przykładów uczących (Tabela 7) należących do:

- klasy 1: „zmiana w  $k + 1$  kroku” = „spadek” (246 przykładów),
- klasy 2: „zmiana w  $k + 1$  kroku” = „bez zmian” (119 przykładów) i
- klasy 3: „zmiana w  $k + 1$  kroku” = „wzrost” (235 przykładów)

tworzono zbiory reduktów jak również zbiory wszystkich minimalnych reguł oraz minimalne zbiory reguł dla rozpatrywanych klas. Tworzone reguły prognozowały zmiany w  $k+1$  kroku.

Poniżej przedstawiono utworzone dla klasy 1 określonej przez warunek („zmiana w  $k + 1$  kroku” = „spadek”) minimalny zbiór reguł, bez relaksacji, ze 100% dopasowaniem (tzn. utworzone reguły opisują poprawnie 100% przykładów uczących należących do klasy 1 i klasy 3),  $q \in [0,1]$  określa stopień przynależności do danej klasy (zależny od liczby przykładów ze zbioru uczącego opisanych przez daną regułę), jeśli część przesłankowa reguły jest spełniona.

- |  |                                   |
|--|-----------------------------------|
| 1) JEŻELI ( $a_{15}=2$ ) TO ( $a_d = „spadek”$ )   | opisuje 101 przykładów, $q=0.210$ |
| 2) JEŻELI ( $a_{14}=2$ ) TO ( $a_d = „spadek”$ )   | opisuje 101 przykładów, $q=0.210$ |
| 3) JEŻELI ( $a_{15}=3$ ) TO ( $a_d = „spadek”$ )   | opisuje 86 przykładów, $q=0.179$  |
| 4) JEŻELI ( $a_{14}=1$ ) TO ( $a_d = „spadek”$ )   | opisuje 41 przykładów, $q=0.085$  |
| 5) JEŻELI ( $a_5=8$ ) TO ( $a_d = „spadek”$ )  | opisuje 10 przykładów, $q=0.021$  |
| 6) JEŻELI ( $a_1=5$ ) $\wedge$ ( $a_{12}=5$ ) $\wedge$ ( $a_{15}=4$ ) TO ( $a_d = „spadek”$ )  | opisuje 10 p. $q=0.021$           |
| 7) JEŻELI ( $a_1=5$ ) $\wedge$ ( $a_2=5$ ) $\wedge$ ( $a_{15}=4$ ) TO ( $a_d = „spadek”$ )     | opisuje 6 p. $q=0.012$            |
| 8) JEŻELI ( $a_2=5$ ) $\wedge$ ( $a_4=4$ ) $\wedge$ ( $a_{15}=4$ ) TO ( $a_d = „spadek”$ )     | opisuje 5 p. $q=0.010$            |
| 9) JEŻELI ( $a_1=5$ ) $\wedge$ ( $a_2=4$ ) $\wedge$ ( $a_{15}=4$ ) TO ( $a_d = „spadek”$ )     | opisuje 3 p. $q=0.006$            |
| 10) JEŻELI ( $a_1=6$ ) $\wedge$ ( $a_{14}=4$ ) $\wedge$ ( $a_{15}=4$ ) TO ( $a_d = „spadek”$ ) | opisuje 2 p. $q=0.004$            |
| 11) JEŻELI ( $a_6=4$ ) $\wedge$ ( $a_{10}=6$ ) $\wedge$ ( $a_{15}=4$ ) TO ( $a_d = „spadek”$ ) | opisuje 2 p. $q=0.004$            |
| 12) JEŻELI ( $a_1=4$ ) $\wedge$ ( $a_8=3$ ) $\wedge$ ( $a_{15}=4$ ) TO ( $a_d = „spadek”$ )    | opisuje 1 p. $q=0.002$            |

Dla klasy 2: („zmiana w  $k + 1$  kroku” = „bez zmian”) utworzono minimalny zbiór 45 reguł, które opisują poprawnie 99% przykładów uczących (opisują 2 przykłady z klasy 1 i 4 przykłady z klasy 3).

Poniżej przedstawiono minimalny zbiór reguł utworzonych dla klasy 3: („zmiana w  $k + 1$  kroku” = „wzrost”), bez relaksacji, ze 100% dopasowaniem (tzn. utworzone reguły opisują poprawnie 100% przykładów uczących należących do klasy 1 i klasy 3).

- |  |                                  |
|--|----------------------------------|
| 1) JEŻELI ( $a_{15}=5$ ) TO ( $a_d = „wzrost”$ ) | opisuje 89 przykładów, $q=0.185$ |
| 2) JEŻELI ( $a_{15}=6$ ) TO ( $a_d = „wzrost”$ ) | opisuje 60 przykładów, $q=0.125$ |
| 3) JEŻELI ( $a_{14}=6$ ) TO ( $a_d = „wzrost”$ ) | opisuje 57 przykładów, $q=0.119$ |
| 4) JEŻELI ( $a_{12}=7$ ) TO ( $a_d = „wzrost”$ ) | opisuje 47 przykładów, $q=0.098$ |

- 5) JEŻELI ( $a_{15}=7$ ) TO ( $a_x = „wzrosł”$ ) opisuje 47 przykładów,  $q=0.098$   
 6) JEŻELI ( $a_{15}=2$ )  $\wedge$  ( $a_{15}=4$ ) TO ( $a_x = „wzrosł”$ ) opisuje 8 przykładów  $q=0.017$   
 7) JEŻELI ( $a_{12}=3$ )  $\wedge$  ( $a_{15}=4$ ) TO ( $a_x = „wzrosł”$ ) opisuje 15 przykładów  $q=0.031$   
 8) JEŻELI ( $a_4=6$ )  $\wedge$  ( $a_{10}=4$ ) TO ( $a_x = „wzrosł”$ ) opisuje 18 przykładów  $q=0.037$   
 9) JEŻELI ( $a_{15}=8$ ) TO ( $a_x = „wzrosł”$ ) opisuje 13 przykładów,  $q=0.027$   
 10) JEŻELI ( $a_4=6$ )  $\wedge$  ( $a_{14}=5$ ) TO ( $a_x = „wzrosł”$ ) opisuje 18 przykładów  $q=0.037$   
 11) JEŻELI ( $a_{11}=6$ )  $\wedge$  ( $a_{14}=5$ ) TO ( $a_x = „wzrosł”$ ) opisuje 9 przykładów  $q=0.019$

W tabeli 8 zamieszczono dla każdej cechy wybranej do opisu przykładów (1 kolumna) tylko te wartości danej cechy, które występują w warunkach tworzących część przesłankową utworzonych reguł dla klasy 1 (2 kolumna) i reguł dla klasy 3 (3 kolumna), wraz ze stopniem przynależności reguły do rozpatrywanej klasy.

Tabela 8. Wartości występujące w regułach dla klasy 1 i klasy 3.

cecha	Wartości występujące w regułach dla klasy 1	Wartości występujące w regułach dla klasy 3
a1	(4; 0.002), (5; 0.040), (6; 0.004)	(2; 0.017)
a2	(4; 0.006), (5; 0.023)	
a3		
a4	(4; 0.010)	(6; 0.075)
a5	(8; 0.021)	
a6	(4; 0.004)	
a7		
a8	(3; 0.002)	
a9		
a10	(6; 0.004)	(4; 0.037)
a11		(6; 0.019)
a12	(5; 0.021)	(3; 0.031), (7; 0.098)
a13		
a14	(1; 0.085), (2; 0.210), (4; 0.004)	(5; 0.056), (6; 0.119)
a15	(2; 0.210), (3; 0.179), (4; 0.060)	(4; 0.048), (5; 0.185), (6; 0.125) (7; 0.098), (8; 0.0270),

W powyższej tabeli zastosowano następujące oznaczenia dla wartości lingwistycznych:

- 1: „duży spadek”, 2: „średni spadek”, 3: „mały spadek”,  
 4: „bez zmian”,  
 5: „niewielki wzrost”, 6: „mały wzrost”, 7: „średni wzrost” i 8: „duży wzrost”.



## 4.2 Dane Nr 2

Dla danych nr 2 przyjęto założenie, że stan  $y(k+1)$  zależy od:

- średniej wartości stanów w ostatnich pięciu, ostatnich czterech, ostatnich trzech, ostatnich dwóch chwilach i od przedostatniej chwili poprzedzającej chwilę  $(k)$ ,
- zmian wartości stanów w ostatnich pięciu chwilach, czterech, trzech i dwóch ostatnich chwilach poprzedzających chwilę  $(k)$ ,
- różnicy pomiędzy maksymalną i minimalną wartością w ostatnich pięciu chwilach, czterech chwilach, trzech chwilach oraz dwóch chwilach poprzedzających chwilę  $(k)$ ,
- stanu  $y(k)$ , tzn.

$$y(k+1) = \Phi \begin{bmatrix} y_{sr}(k-5), y_{sr}(k-4), y_{sr}(k-3), y_{sr}(k-2), y_{sr}(k-1), \\ lzm(k-5), lzm(k-4), lzm(k-3), lzm(k-2), \\ mm(k-5), mm(k-4), mm(k-3), mm(k-2) \\ y(k), \end{bmatrix} \quad (12)$$

gdzie funkcja  $\Phi$  jest nieznaną funkcją odwzorującą wyżej wymienione czynniki w prognozowany stan w chwili  $(k+1)$ . Przyjęto następujące zależności:

$$y_{sr}(k-5) = \sum_{i=k-5}^{k-1} [x(i)] / 5$$

$$y_{sr}(k-4) = \sum_{i=k-4}^{k-1} [x(i)] / 4$$

$$y_{sr}(k-3) = \sum_{i=k-3}^{k-1} [x(i)] / 3$$

$$y_{sr}(k-2) = \sum_{i=k-2}^{k-1} [x(i)] / 2$$

$$y_{sr}(k-1) = x(k-1)$$

$$lzm(k-5) = x(k-5) / x(k-1) / 5$$

$$lzm(k-4) = x(k-4) / x(k-1)$$

$$lzm(k-3) = x(k-3) / x(k-1) / 3$$

$$lzm(k-2) = x(k-2) / x(k-1) / 2$$

$$mm(k-5) = \left[ \begin{array}{l} \max\{x(k-5), x(k-4), x(k-3), x(k-2), x(k-1)\} \\ - \min\{x(k-5), x(k-4), x(k-3), x(k-2), x(k-1)\} \end{array} \right] / 5$$

$$mm(k-4) = \left[ \begin{array}{l} \max\{s(k-4), s(k-3), s(k-2), s(k-1)\} \\ - \min\{s(k-4), s(k-3), s(k-2), s(k-1)\} \end{array} \right] / 4$$

$$mm(k-3) = \left[ \begin{array}{l} \max\{x(k-3), x(k-2), x(k-1)\} \\ - \min\{x(k-3), x(k-2), x(k-1)\} \end{array} \right] / 3$$

$$mm(k-2) = \left[ \begin{array}{l} \max\{x(k-2), x(k-1)\} \\ - \min\{x(k-2), x(k-1)\} \end{array} \right] / 2$$

$$y(k) = x(k)$$

$$y(k+1) = x(k+1)$$

W Tabeli 9 przedstawiono wartości cech pierwotnych, tzn. dla wartości sygnałów.

Tabela 9.

$k$	$y_{sr}(k-5)$	$y_{sr}(k-4)$	$y_{sr}(k-3)$	$y_{sr}(k-2)$	$y_{sr}(k-1)$
$\mathcal{X}$	1,1726	1,1831	1,1928	1,2018	1,2102
$\mathcal{X}+1$	1,1915	1,2009	1,2096	1,2178	1,2253
$\mathcal{X}+2$	1,2085	1,2170	1,2248	1,2321	1,2390
$\mathcal{X}+3$	1,2238	1,2314	1,2385	1,2451	1,2512
$\mathcal{X}+4$	1,2376	1,2444	1,2508	1,2567	1,2622
$\mathcal{X}+5$	1,2500	1,2561	1,2619	1,2672	1,2722
$\vdots$					

$k$	$lzm(k-5)$	$lzm(k-4)$	$lzm(k-3)$	$lzm(k-2)$
$\mathcal{X}$	0,0000	0,0000	0,0000	0,0000
$\mathcal{X}+1$	0,0000	0,0000	0,0000	0,0000
$\mathcal{X}+2$	0,0000	0,0000	0,0000	0,0000
$\mathcal{X}+3$	0,0000	0,0000	0,0000	0,0000
$\mathcal{X}+4$	0,0000	0,0000	0,0000	0,0000
$\mathcal{X}+5$	0,0000	0,0000	0,0000	0,0000
$\vdots$				

$k$	$mm(k-5)$	$mm(k-4)$	$mm(k-3)$	$mm(k-2)$	$y(k)$	$y(k+1)$
$\mathcal{X}$	0,1587	0,1406	0,1183	0,0841	1,2102	1,2253
$\mathcal{X}+1$	0,1428	0,1266	0,1065	0,0757	1,2253	1,2390
$\mathcal{X}+2$	0,1285	0,1139	0,0958	0,0681	1,2390	1,2512
$\mathcal{X}+3$	0,1157	0,1025	0,0863	0,0613	1,2512	1,2622
$\mathcal{X}+4$	0,1041	0,0923	0,0776	0,0552	1,2622	1,2722
$\mathcal{X}+5$	0,0937	0,0831	0,0699	0,0496	1,2722	1,2882
$\vdots$						

Przejdźcie na właściwe przykłady uczące,  $\Delta x(i) = x(i) - x(i-1)$

$$a1 = \sum_{i=k-5}^{k-1} \Delta[x(i)] \times 100/5 + 4$$

$$a2 = \sum_{i=k-4}^{k-1} [\Delta x(i)] \times 100/4 + 4$$

$$a3 = \sum_{i=k-3}^{k-1} [\Delta x(i)] \times 100/3 + 4$$

$$a4 = \sum_{i=k-2}^{k-1} [\Delta x(i)] \times 100/2 + 4$$

$$a5 = \Delta x(k-1) \times 100 + 4$$

$$a6 = \{\Delta x(k-5) / \Delta x(k-1)\} \times 10/5$$

$$a7 = \{\Delta x(k-4) / \Delta x(k-1)\} \times 10/4$$

$$a8 = \{\Delta x(k-3) / \Delta x(k-1)\} \times 10/3$$

$$a9 = \{\Delta x(k-2) / \Delta x(k-1)\} \times 10/2$$

$$a10 = \{\Delta x(k-1)\} \times 10/2$$

$$a11 = \left[ \begin{array}{l} \max\{\Delta x(k-5), \Delta x(k-4), \Delta x(k-3), \Delta x(k-2), \Delta x(k-1)\} \\ - \min\{\Delta x(k-5), \Delta x(k-4), \Delta x(k-3), \Delta x(k-2), \Delta x(k-1)\} \end{array} \right] \times 100/5 + 4$$

$$a12 = \left[ \begin{array}{l} \max\{\Delta x(k-4), \Delta x(k-3), \Delta x(k-2), \Delta x(k-1)\} \\ - \min\{\Delta x(k-4), \Delta x(k-3), \Delta x(k-2), \Delta x(k-1)\} \end{array} \right] \times 100/4 + 4$$

$$a13 = \left[ \begin{array}{l} \max\{\Delta x(k-3), \Delta x(k-2), \Delta x(k-1)\} \\ - \min\{\Delta x(k-3), \Delta x(k-2), \Delta x(k-1)\} \end{array} \right] \times 100/3 + 4$$

$$a14 = \left[ \begin{array}{l} \max\{\Delta x(k-2), \Delta x(k-1)\} \\ - \min\{\Delta x(k-2), \Delta x(k-1)\} \end{array} \right] \times 100/2 + 4$$

$$a15 = y(k) = x(k)$$

Cechę decyzyjną określono w postaci:  $a_j$ : „zmiana w  $k+1$  kroku” z wartościami lingwistycznymi: 1: „spadek”, 2: „bez zmian”, 3: „wzrost”.

$$a_j = y(k+1) = x(k+1)$$

W Tabeli 11 zamieszczono utworzone przykłady uczące

Tabela 11.

Lp.	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	a15	$a_j$
u1	5	5	5	5	5	0	0	0	0	0	1	1	1	0	5	3
u2	5	5	5	5	5	0	0	0	0	0	1	1	1	0	5	3
u3	5	5	5	5	5	0	0	0	0	0	1	1	1	0	5	3
u4	5	5	5	5	5	0	0	0	0	0	1	1	0	0	5	3
u5	5	5	5	5	5	0	0	0	0	0	1	1	0	0	5	2
u6	5	5	5	5	4	0	0	0	0	0	1	0	0	0	4	3

u7	5	5	5	5	5	0	0	0	0	0	1	1	1	2	5	3
u8	5	5	5	5	5	0	0	0	0	0	1	1	1	0	5	2
u9	5	5	5	5	4	0	0	0	0	0	1	1	1	2	4	2
u10	4	4	4	4	4	2	3	3	5	10	2	3	3	3	4	1
u11	4	4	4	3	3	2	3	3	5	0	4	4	4	4	3	1
u12	4	3	3	3	2	2	3	3	0	0	5	5	5	3	2	1
u13	3	3	2	2	2	2	3	0	0	0	5	5	4	3	2	1
u14	2	2	2	2	1	2	0	0	0	0	5	4	3	2	1	1
u15	2	2	1	1	1	0	0	0	0	0	4	3	2	1	1	1
u16	2	1	1	1	1	0	0	0	0	0	3	2	1	1	1	1
u17	1	1	1	1	1	0	0	0	0	0	2	1	1	0	1	1
u18	1	1	1	1	1	0	0	0	0	0	1	0	0	0	1	1
u19	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1
u20	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1
u21	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1
u22	1	1	1	1	2	0	0	0	0	0	1	1	1	1	2	1
u23	1	1	2	2	2	0	0	0	0	0	1	1	1	1	2	1
u24	2	2	2	2	2	0	0	0	0	0	1	1	1	1	2	1
u25	2	2	2	2	2	0	0	0	0	0	1	1	1	1	2	1
⋮																

Na podstawie zbioru przykładów uczących (Tabela 11) należących do:

- klasy 1: „zmiana w  $k + 1$  kroku” = „spadek” (246 przykładów),  
klasy 2: „zmiana w  $k + 1$  kroku” = „bez zmian” (119 przykładów) i  
klasy 3: „zmiana w  $k + 1$  kroku” = „wzrost” (238 przykładów)

tworzono zbiory reduktów jak również zbiory wszystkich minimalnych reguł oraz minimalne zbiory reguł dla rozpatrywanych klas. Tworzone reguły prognozowały zmiany w  $k+1$  kroku.

Poniżej przedstawiono utworzony dla klasy 1 określonej przez warunek („zmiana w  $k + 1$  kroku” = „spadek”) minimalny zbiór reguł, bez relaksacji, ze 100% dopasowaniem (tzn. utworzone reguły opisują poprawnie 100% przykładów uczących należących do klasy 1 i klasy 3),  $q \in [0,1]$  określa stopień przynależności do danej klasy (zależny od liczby przykładów ze zbioru uczącego opisywanych przez daną regułę), jeśli część przesłankowa reguły jest spełniona.

- 1) JEŻELI ( $a_{15}=2$ ) TO ( $a_j = \text{„spadek”}$ ) opisuje 101 przykładów,  $q=0.209$
- 2) JEŻELI ( $a_{15}=3$ ) TO ( $a_j = \text{„spadek”}$ ) opisuje 86 przykładów,  $q=0.179$
- 3) JEŻELI ( $a_{15}=1$ ) TO ( $a_j = \text{„spadek”}$ ) opisuje 41 przykładów,  $q=0.085$
- 4) JEŻELI ( $a_1=4$ )  $\wedge$  ( $a_{14}=3$ ) TO ( $a_j = \text{„spadek”}$ ) opisuje 12 przykładów  $q=0.025$
- 5) JEŻELI ( $a_5=8$ ) TO ( $a_j = \text{„spadek”}$ ) opisuje 10 przykładów,  $q=0.021$
- 6) JEŻELI ( $a_2=5$ )  $\wedge$  ( $a_{10}=10$ ) TO ( $a_j = \text{„spadek”}$ ) opisuje 7 przykładów  $q=0.014$
- 7) JEŻELI ( $a_1=5$ )  $\wedge$  ( $a_2=4$ )  $\wedge$  ( $a_{15}=4$ ) TO ( $a_j = \text{„spadek”}$ ) opisuje 3 p.  $q=0.006$
- 8) JEŻELI ( $a_1=6$ )  $\wedge$  ( $a_{10}=10$ ) TO ( $a_j = \text{„spadek”}$ ) opisuje 2 przykłady  $q=0.004$
- 9) JEŻELI ( $a_{11}=5$ )  $\wedge$  ( $a_{15}=4$ ) TO ( $a_j = \text{„spadek”}$ ) opisuje 2 przykłady  $q=0.004$
- 10) JEŻELI ( $a_6=4$ )  $\wedge$  ( $a_{10}=10$ ) TO ( $a_j = \text{„spadek”}$ ) opisuje 1 przykład  $q=0.002$

Poniżej przedstawiono minimalny zbiór reguł utworzonych dla klasy 3: („zmiana w  $k + 1$  kroku” = „wzrost”), bez relaksacji, które opisują poprawnie ok. 96% przykładów uczących.

- 1) JEŻELI ( $a15=5$ ) TO ( $a_d = „wzrost”$ ) opisuje 92 przykładów,  $q=0.190$
- 2) JEŻELI ( $a15=6$ ) TO ( $a_d = „wzrost”$ ) opisuje 60 przykładów,  $q=0.124$
- 3) JEŻELI ( $a15=7$ ) TO ( $a_d = „wzrost”$ ) opisuje 47 przykładów,  $q=0.097$
- 4) JEŻELI ( $a1=7$ )  $\wedge$  ( $a3=6$ ) TO ( $a_d = „wzrost”$ ) opisuje 19 przykładów  $q=0.039$
- 5) JEŻELI ( $a15=8$ ) TO ( $a_d = „wzrost”$ ) opisuje 13 przykładów,  $q=0.027$
- 6) JEŻELI ( $a4=6$ )  $\wedge$  ( $a15=4$ ) TO ( $a_d = „wzrost”$ ) opisuje 11 przykładów  $q=0.023$
- 7) JEŻELI ( $a1=2$ )  $\wedge$  ( $a13=4$ ) TO ( $a_d = „wzrost”$ ) opisuje 11 przykładów  $q=0.023$
- 8) JEŻELI ( $a5=2$ )  $\wedge$  ( $a12=4$ ) TO ( $a_d = „wzrost”$ ) opisuje 6 przykładów  $q=0.012$
- 9) JEŻELI ( $a3=2$ )  $\wedge$  ( $a12=4$ )  $\wedge$  ( $a15=4$ ) TO ( $a_d = „wzrost”$ ) opisuje 3 p.  $q=0.006$
- 10) JEŻELI ( $a1=1$ )  $\wedge$  ( $a13=4$ ) TO ( $a_d = „wzrost”$ ) opisuje 3 przykłady  $q=0.006$
- 11) JEŻELI ( $a12=0$ )  $\wedge$  ( $a15=4$ ) TO ( $a_d = „wzrost”$ ) opisuje 2 przykłady  $q=0.004$
- 12) JEŻELI ( $a14=10$ )  $\wedge$  ( $a15=4$ ) TO ( $a_d = „wzrost”$ ) opisuje 2 przykłady  $q=0.004$
- 13) JEŻELI ( $a12=4$ )  $\wedge$  ( $a13=3$ )  $\wedge$  ( $a15=4$ ) TO ( $a_d = „wzrost”$ ) opisuje 1 p.  $q=0.002$
- 14) JEŻELI ( $a1=1$ )  $\wedge$  ( $a13=3$ ) TO ( $a_d = „wzrost”$ ) opisuje 1 przykład  $q=0.002$

Tabela 12. Wartości występujące w regułach dla klasy 1 i klasy 3.

cecha	Wartości występujące w regułach dla klasy 1	Wartości występujące w regułach dla klasy 3
a1	(4; 0.025), (5; 0.006), (6; 0.004)	(1; 0.008), (2; 0.023), (7; 0.039)
a2	(4; 0.006), (5; 0.014)	
a3		(2; 0.006), (1; 0.039),
a4		(6; 0.023)
a5	(8; 0.021)	(2; 0.012),
a6	(4; 0.002)	
a7		
a8		
a9		
a10	(10; 0.021)	
a11	(5; 0.004)	
a12		(4; 0.021)
a13		(3; 0.004), (4; 0.029)
a14	(3; 0.025)	
a15	(1; 0.085), (2; 0.209), (3; 0.178), (4; 0.010)	(4; 0.039), (5; 0.190), (6; 0.124) (7; 0.097), (8; 0.027),

W tabeli 12 zamieszczono dla każdej cechy wybranej do opisu przykładów (1 kolumna) tylko te wartości danej cechy, które występują w warunkach tworzących część przesłankową utworzonych reguł dla klasy 1 (2 kolumna) i reguł dla klasy 3 (3 kolumna), wraz ze stopniem przynależności reguły do rozpatrywanej klasy.

## 5 Uwagi końcowe

W pracy podjęto próbę przedstawienia zadania predykcji szeregu czasowego jako zadania uczenia maszynowego na podstawie przykładów i opisanie problemu predykcji w postaci zbioru reguł decyzyjnych.

Wybrane metody uczenia maszynowego na podstawie przykładów zastosowano do predykcji szeregu czasowego wygenerowanego w oparciu o funkcję Mackey'a-Glassa. Analizowano klasy określone przez warunki: („zmiana w  $k + 1$  kroku” = „spadek”), („zmiana w  $k + 1$  kroku” = „bez zmian”) i („zmiana w  $k + 1$  kroku” = „wzrost”). Zaproponowano wstępny sposób obróbki danych wejściowych, tzn. wybór cech, dyskretyzację wartości tych cech oraz ich zapis w postaci lingwistycznej.

Utworzone minimalne zbiory reguł zawierają małą liczbę prostych reguł, z małą liczbą warunków je tworzących. Uzyskano bardzo dobre dopasowanie reguł do danych uczących i dużą dokładność klasyfikacji dla danych testowych.

Planowana jest rozszerzona analiza innych cech, analiza większej liczby klas oraz przeprowadzenie obliczeń dla innych szeregów czasowych.

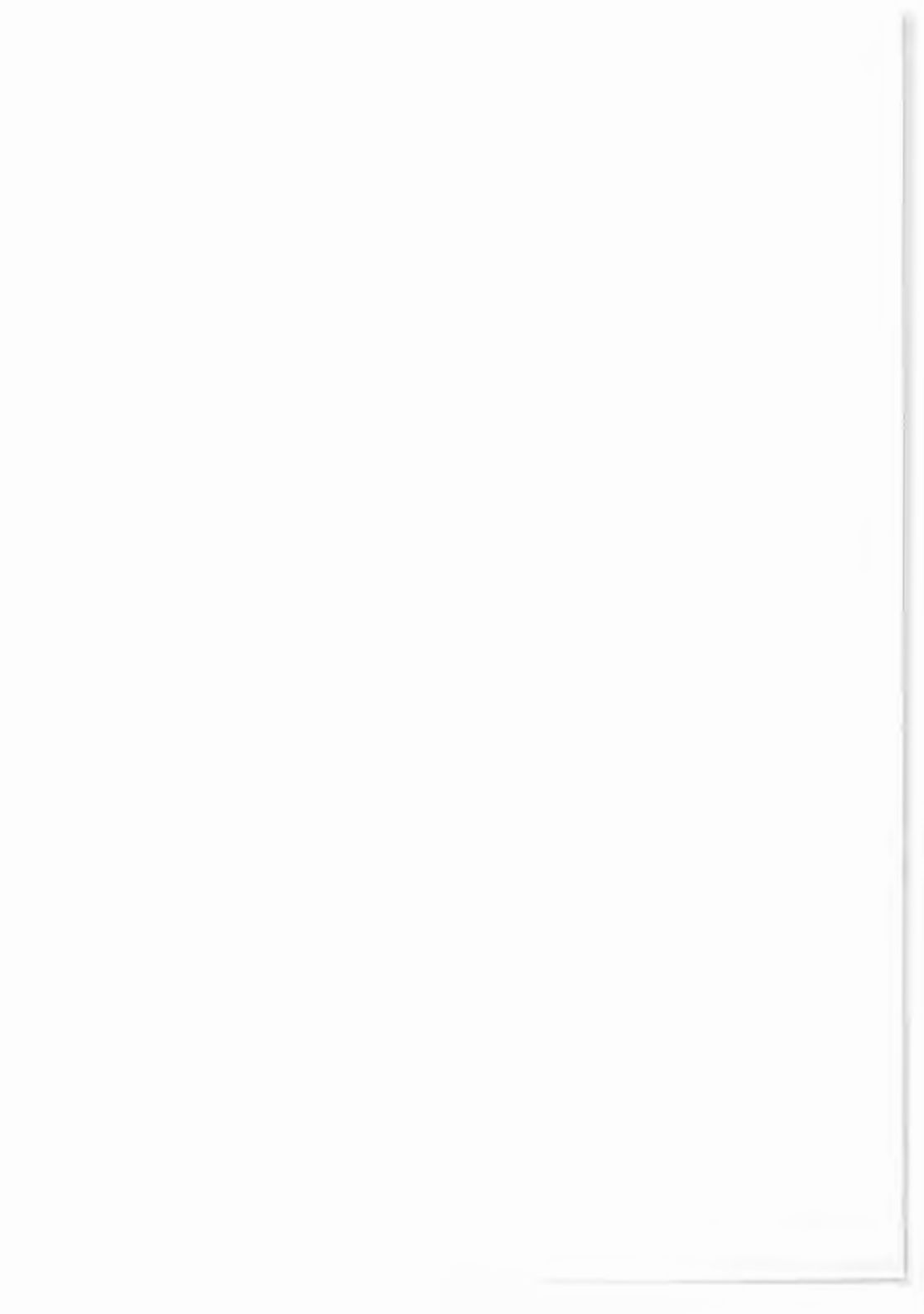
## Literatura

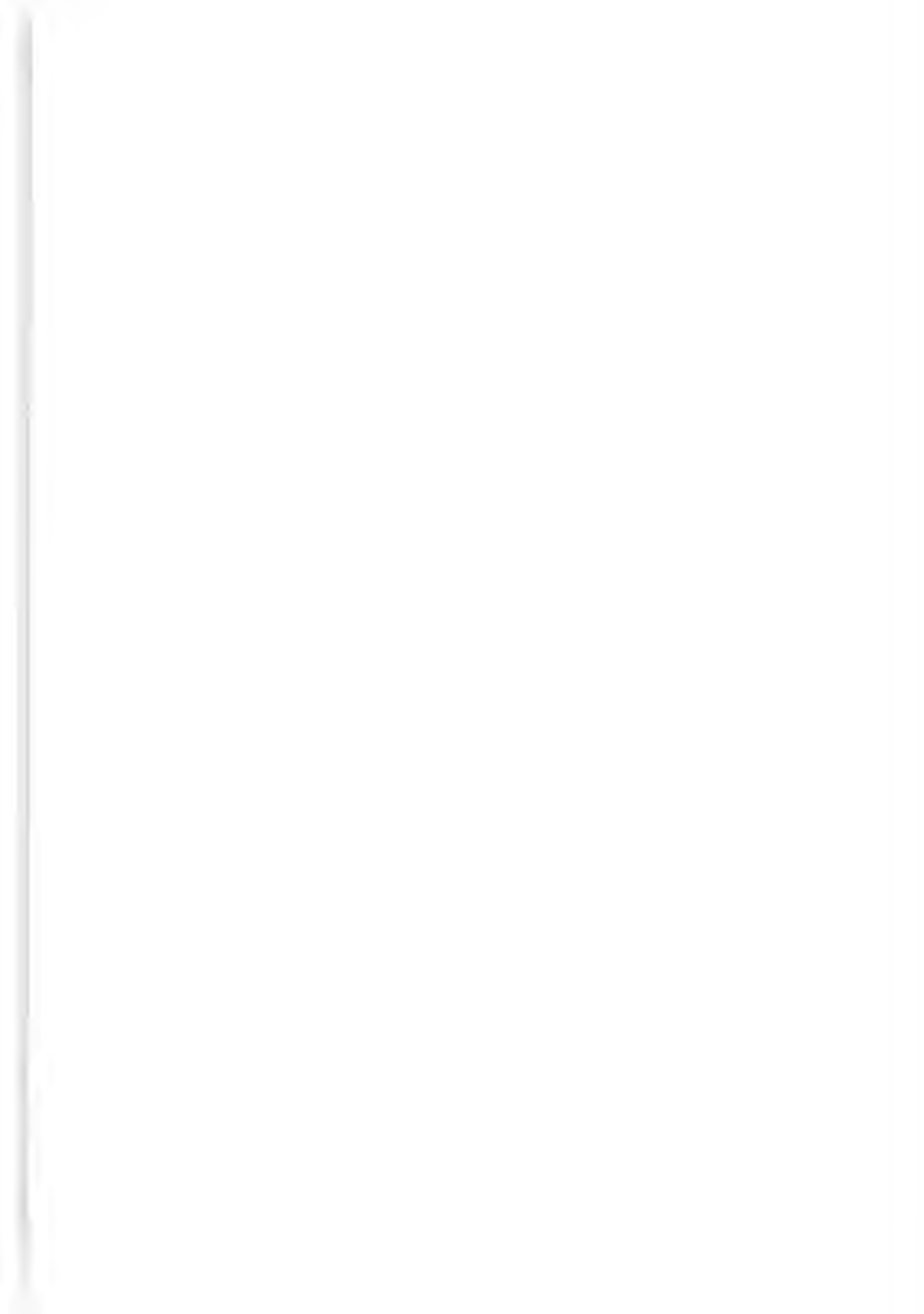
1. Kacprzyk J., Szkatuła G. (1996) An algorithm for learning from erroneous and incorrigible examples, *Int. J. of Intelligent Syst.* 11, 565 - 582.
2. Kacprzyk J., Szkatuła G. (1997) An improved inductive learning algorithm with a preanalysis of data", In: Z.W. Raś and A. Skowron (eds.): *Foundations of Intelligent Systems* (Proceedings of 10th ISMIS'97 Symposium, Charlotte, uSA), Springer-Verlag, Berlin, 157 - 166.
3. Kacprzyk J., Szkatuła G. (1998) *IP1 - An Improved Inductive Learning Procedure with a Preprocessing of Data*. Proceedings of IDEAL'98 (Hong Kong), Springer-Verlag.
4. Kacprzyk J., Szkatuła G. (1999) An inductive learning algorithm with a preanalysis of data. *International Journal of Knowledge - Based Intelligent Engineering Systems*, vol. 3, 135-146.
5. Kacprzyk J., Szkatuła G. (2002) An integer programming approach to inductive learning using genetic and greedy algorithms. W: L.C. Jain and J.Kacprzyk (eds.) *New learning paradigms in soft computing*. Studies in Fuzziness and Soft Computing. Physica-Verlag Heidelberg, 323 - 367.
6. Krawczak, M., Miklewski, A., Jakubowski, A., Konieczny, P. (2000) Investment Risk Management, (in Polish). Polish Academy of Sciences, Systems Research 25.

7. Kunchewa L.I. (2000) *Fuzzy Classifier Design*. Physica-Verlag.
8. Pawlak Z. (1982): Rough Set. International Journal of Computer and Information Sciences, Vol. 11, No 5, 341-356
9. Pawlak Z. (1991): Rough Set. Theoretical Aspect of Reasoning about Data. Kluwer Academic Publishers.
10. Shaw, R. S. 1981. "Strange Attractors, Chaotic Behavior and Information Flow." Z. Naturforsch. 36A: 80-112.
11. Schuster, H.G. 1995. Chaos deterministyczny. PWN. Warszawa.
12. Szkatuła G. (1995) Machine learning from examples under errors in data, Ph.D. thesis, SRI PAS Warsaw, Poland.
13. Szkatuła G. (2002). *Zastosowanie zmodyfikowanego zadania pokrycia w uczeniu maszynowym*. W: Gutenbaum J. (eds.): Automatyka Sterowanie Zarządzanie. SRI PAS, Warszawa, 431-445.
14. Weigend, A. S., B. A. Huberman, and D. E. Rumelhart. 1990. Predicting the Future: A Connectionist Approach." International Journal of Neural Systems 1:193-209.
15. Weigend, A., B. A. Huberman, and D. E. Rumelhart. 1992. "Predicting Sunspots and Exchange Rates with Connectionist Networks." In Nonlinear Modeling and Forecasting, edited by M. Casdagli and S. Eubank, 395-432. Redwood City, CA: Addison-Wesley.
16. Weigend, A. S., and B. LeBaron. 1994. "Evaluating Neural Network Predictors by Bootstrapping." In Proceedings of International Conference on Neural Information Processing (ICONIP'94), Seoul, Korea. Also Technical Report Cu-CS-725-94, Computer Science Department, university of Colorado at Boulder.
17. Yule G. (1927) On a Method of Investigating Periodicity in Disturbed Series with Special Reference to Wolfer's Sunspot Numbers. Phil. Trans. Roy. Soc. London A 226: 267-298.











the 1990s, the number of people in the UK who are aged 65 and over has increased from 10.5 million to 13.5 million (13.5% of the population).

There are a number of reasons why the number of people aged 65 and over has increased. One of the main reasons is that people are living longer. The life expectancy at birth in the UK is now 77 years for men and 81 years for women. This is a significant increase from the 1950s, when life expectancy at birth was 71 years for men and 75 years for women.

Another reason why the number of people aged 65 and over has increased is that people are staying in the workforce longer. In the 1950s, most people retired at the age of 65. Now, many people continue to work until they are 70 or even 75 years old.

There are a number of reasons why people are staying in the workforce longer. One of the main reasons is that people are healthier. In the 1950s, many people were in poor health and were unable to work. Now, people are living longer and healthier lives, so they are able to work for longer.

Another reason why people are staying in the workforce longer is that they need the money. In the 1950s, most people had a pension. Now, many people do not have a pension and need to work to support themselves.

There are a number of reasons why people need the money. One of the main reasons is that the cost of living has increased. In the 1950s, the cost of living was much lower than it is now. Now, people need more money to live on.

Another reason why people need the money is that they have a large family. In the 1950s, most people had a small family. Now, many people have a large family and need more money to support them.

There are a number of reasons why people have a large family. One of the main reasons is that people are having children later in life. In the 1950s, most people had their first child at the age of 20. Now, many people have their first child at the age of 30 or even 40.

Another reason why people have a large family is that they are having more children. In the 1950s, most people had two or three children. Now, many people have four or more children.

There are a number of reasons why people are having more children. One of the main reasons is that people are having children for longer. In the 1950s, most people stopped having children at the age of 35. Now, many people continue to have children until they are 40 or even 45.

Another reason why people are having more children is that they are having children who are healthier. In the 1950s, many children were born with health problems. Now, children are living longer and healthier lives, so people are having more children.

There are a number of reasons why children are living longer and healthier lives. One of the main reasons is that people are having children who are better educated. In the 1950s, most children did not go to university. Now, many children go to university and are better educated than their parents.

Another reason why children are living longer and healthier lives is that they are having children who are better cared for. In the 1950s, many children were not well cared for. Now, children are better cared for and are living longer and healthier lives.