

Agnieszka GORNIG, Antoni WOJACZEK
Politechnika Śląska, Gliwice

WYZNACZANIE RZECZYWISTYCH KLAS WĘGLA W OPARCIU O ALGORYTM ROZMYTEJ KLASYFIKACJI DANYCH

Streszczenie. W referacie przedstawiono przykład wyznaczania rzeczywistych klas węgla w oparciu o algorytm rozmytego grupowania danych. Klasyfikacji poddano zbiór zawierający pomiary próbek węgla w KWK „Knurów” i wyznaczono 5 klas reprezentowanych przez współrzędne centroidów. Dla każdej klasy postawiono prognozę ceny węgla według formuł Blaschkego w oparciu o ceny węgla wskaźnikowego oraz ceny węgla wzorcowego. Wykorzystując rozkład zmian klas węgla w czasie wyznaczono wagi reguł oraz utworzono bazę reguł będącą rozmytym modelem typu Mamdaniego.

REAL COAL CLUSTERS DELIMITATION BASED ON THE ALGORITHM OF FUZZY DATE CLASSIFICATION

Summary. In the article an example of the real coal clusters delimitation is presented. The algorithm of fuzzy date classification is used to make partitions of date. A file of 500 coal samples is divided. The program creates a required number of clusters and owns the fuzzy membership coefficients for the samples. A procedure of auto-correction assures optimal classification.

1. Wprowadzenie

O jakości węgla decyduje przede wszystkim zawartość węgla i popiołu, oraz jakość właściwej masy węglowej, której najważniejszym wykładnikiem jest stopień uwęglenia. Wszystkie klasyfikacje węgla opierające się na jego wskaźnikach chemicznych dzielą węgiel według stopnia uwęglenia. Węgiel kamienny charakteryzuje stopień uwęglenia rzędu (76 - 94%).

Jakość węgla charakteryzuje się parametrami jakościowymi, np. wartością opałową, zawartością: siarki, popiołu, wilgoci. Są to parametry cenotwórcze.

Węgiel surowy składa się z różnej wielkości ziarn, z których każde ma określoną gęstość w przybliżeniu proporcjonalną do zawartości popiołu. Zależności pomiędzy gęstością ziarna a parametrami jakościowymi przedstawia się za pomocą krzywych wzbogacalności. Do celów analizy densymetrycznej poszczególne ziarna węglowe grupuje się w pewne frakcje ciężarowe o określonej dolnej i górnej gęstości oraz oblicza się udział każdej z frakcji w urobku surowym. W procesach wzbogacania węgla dobiera się gęstości rozdziału, reguluje parametry jakościowe produktów końcowych zgodnie z przebiegiem krzywych charakterystyk technologicznych. Wobec tego jakość węgla zmienia się w trakcie procesów wzbogacania stosujących różne gęstości rozdziału pomiędzy koncentratem, odpadami i produktem pośrednim.

Zmienna jakość węgla wpływa na różne wartości gospodarcze węgla. Decyduje to o złożoności wyznaczania poprawnych relacji pomiędzy jakością węgla a jego ceną. Dlatego system cen oparty na właściwościach technologicznych węgla powinien uwzględniać indywidualne cechy każdego urobku. Dla producentów węgla korzystniejsze jest funkcjonowanie indywidualnych cenników węglowych dostosowanych do oferty jakościowej węgla tego zakładu. Stworzenie takich cenników wymaga znajomości klas węgla, które powinny jak najwierniej przedstawiać właściwości węgla. Rozwiązaniem może być wyznaczenie rzeczywistych klas węgla w oparciu o istniejący zbiór właściwości węgla.

2. Rozmyta klasyfikacja danych

Algorytm grupowania próbek węgla wokół c -centroidów jest przykładem rozmytej analizy grupowania danych typu *clustering*. Analiza ta polega na poszukiwaniu grup zwanych *klastrami* wśród danych. Dąży do tworzenia grup, w których obiekty należące do tej samej grupy są maksymalnie podobne do siebie, podczas gdy obiekty w różnych grupach są maksymalnie różne. Podobieństwo danych określa się na podstawie przyjętej miary podobieństwa oraz kryterium oceny tej miary.

Niech zbiór danych $X = \{x_1, \dots, x_p\} \subset R^p$ jest podzbiorem rzeczywistej R^p przestrzeni p -wymiarowych wektorów. Każdy $x_k = (x_{k1}, \dots, x_{kp}) \in R^p$ jest wektorem właściwości

opisującym pojedynczy obiekt zbioru danych. X_{ki} jest i -tą właściwością obiektu x_k . Niepodobieństwo dwóch obiektów zwane **odległością** między dwoma obiektami x_k oraz x_l jest funkcją o wartościach ze zbioru liczb rzeczywistych $d: X \times X \rightarrow R^+$.

W ujęciu rozmytym każda dana należy do każdego klastra, ale w różnym stopniu. Z każdym klastrem można powiązać pewien obszar decyzyjny. Każdy podział zbioru $X = \{x_1, \dots, x_n\}$ na podzbiory \tilde{S}_i ($i = 1, \dots, c$) rozmyte można opisać przez funkcję przynależności $\mu_{\tilde{S}_i}$. Funkcja ta spełnia warunek:

$$\mu_{\tilde{S}_i} : X \rightarrow [0, 1] \quad (2.1)$$

gdzie $\mu_{\tilde{S}_i}$ określają stopień przynależności obiektu x_k do podzbioru \tilde{S}_i :

$$\mu_{ik} := \mu_{\tilde{S}_i}(x_k) \quad (2.2)$$

Niech V_{cn} jest zbiorem wszystkich rzeczywistych macierzy $c \times n$ oraz $2 \leq c < n$ jest liczbą całkowitą. $\tilde{U} = [\mu_{ik}] \in V_{cn}$ jest zwane **rozmytym c -podziałem**, jeżeli spełnia następujące warunki:

$$1. \mu_{ik} \in [0, 1] \quad 1 \leq i \leq c, 1 \leq k \leq n \quad (2.3)$$

$$2. \sum_{i=1}^c \mu_{ik} = 1 \quad 1 \leq k \leq n \quad (2.4)$$

$$3. 0 < \sum_{k=1}^n \mu_{ik} < n \quad 1 \leq i \leq c \quad (2.5)$$

Położenie klastrów określa „środek klastra” zwany **centroidem**: $v_i = (v_{i1}, \dots, v_{ip}) \in R^p$, $i = 1, \dots, c$ wokół którego leżą obiekty zgrupowane w tym klastrze.

Wykorzystany w badaniach algorytm rozmytej klasyfikacji danych wyznacza kryterium podziału za pomocą funkcji celu. Lokalne ekstremum tej funkcji wskazuje optymalny podział danych. Jest ono wyznaczane w oparciu o oszacowywanie podobieństwa kolejnych podziałów dla każdej liczby klastrów c według wybranej funkcji celu.

Do optymalizacji wstępnego podziału wykorzystano kryterium wariancji dla podziału rozmytego. Mierzy ono niepodobieństwo między każdym punktem w klastrze a środkiem klastra w sensie odległości euklidesowej obliczanej jako:

$$d_{ik} = d(x_k, v_i) = \|x_k - v_i\| = \left[\sum_{j=1}^p (x_{kj} - v_{ij})^2 \right]^{1/2} \quad (2.6)$$

Kryterium wariancji odpowiada minimalizacji sumy wariancji wszystkich zmiennych j w każdym klastrze i . Sprowadza problem optymalizacji grupowania do rozwiązania następującego problemu:

$$\min z(U, v) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m \|x_k - v_i\|^2 \quad (2.7)$$

gdzie $\tilde{U} \in M_{fc}$ oraz $v \in R^{cp}$ oraz

$$v_i = \frac{1}{\left[\sum_{k=1}^n (\mu_{ik})^m \right]} \sum_{k=1}^n (\mu_{ik})^m x_k, \quad m > 1, \quad i = 1, \dots, c \quad (2.8)$$

gdzie

$$\mu_{ik} = \frac{\left(\frac{1}{\|x_k - v_i\|_G^2} \right)^{1/(m-1)}}{\sum_{j=1}^c \left(\frac{1}{\|x_k - v_j\|_G^2} \right)^{1/(m-1)}}, \quad i = 1, \dots, c; \quad k = 1, \dots, n \quad (2.9)$$

Nie można rozwiązać podanych wyżej równań metodą analityczną. Dlatego wykorzystuje się algorytm iteracyjny podający jako rozwiązanie kolejne przybliżenia funkcji kryterialnej, np. algorytm grupowania rozmytego wokół c -centroidów [1].

3. Grupowanie danych z zastosowaniem algorytmu rozmytej klasyfikacji danych

Analizowany zbiór danych obejmuje wielowymiarowe dane ilościowe zawierające informacje o zawartości frakcji ziaren węgla o gęstości określonej przedziałami oraz o procentowej zawartości popiołu w tych frakcjach. Wartości pochodzą z bazy danych KWK „Knurów”. Kolejne zmienne określają:

- x_1 - zawartość (udział wyrażony ułamkiem) frakcji ziaren węgla o gęstości $\rho < 1,3 \text{ g/cm}^3$,
- x_2 - zawartość frakcji ziaren węgla o gęstości $1,3 < \rho < 1,5 \text{ g/cm}^3$,
- x_3 - zawartość frakcji ziaren węgla o gęstości $1,5 < \rho < 1,8 \text{ g/cm}^3$,
- x_4 - zawartość frakcji ziaren węgla o gęstości $\rho > 1,8 \text{ g/cm}^3$,
- x_5 - zawartość popiołu we frakcji pierwszej o gęstości $\rho < 1,3 \text{ g/cm}^3$,
- x_6 - zawartość popiołu we frakcji drugiej o gęstości $1,3 < \rho < 1,5 \text{ g/cm}^3$,
- x_7 - zawartość popiołu we frakcji trzeciej o gęstości $1,5 < \rho < 1,8 \text{ g/cm}^3$,
- x_8 - zawartość popiołu we frakcji czwartej o gęstości $\rho > 1,8 \text{ g/cm}^3$.

Zadanie klasyfikacji polega na podziale 500 obiektów $x \in X$ znajdujących się w bazie danych scharakteryzowanych przez 8 zmiennych (właściwości) w pięciu jednoznacznie określonych jednorodnych klastrach.

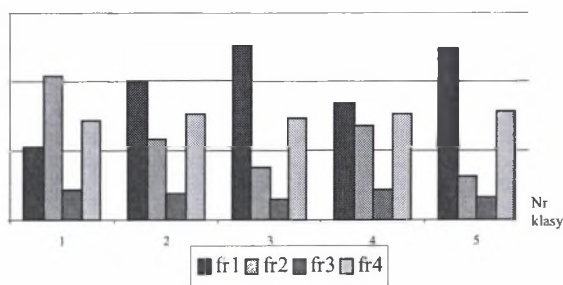
W analizie wykorzystano zmodyfikowany algorytm grupowania wokół c-centroidów, który był przedstawiony na Konferencji Zrównoważonego Rozwoju 2003 w Gliwicach [8]. Jako wynik uzyskano 5 klas, które są scharakteryzowane poprzez wyznaczone wartości średnie zmiennych opisujących próbki węgla znajdujące się w poszczególnych klasach. Otrzymane współrzędne centroidów przedstawia tablica 1 oraz wykresy słupkowe. Rysunek 1 prezentuje udziały frakcji o określonych gęstościach w poszczególnych klasach, a rysunek 2 prezentuje procentowe zawartości popiołu we frakcjach w poszczególnych klasach.

Tablica 1

Wartości zmiennych centroidów

	x1	x2	x3	x4	x5	x6	x7
klasa 1	0,2135	0,4164	0,0846	0,2856	2,3712	5,9538	28,360
klasa 2	0,3929	0,2312	0,0744	0,3016	2,2453	7,7927	31,288
klasa 3	0,5018	0,1495	0,0577	0,291	2,6098	12,0129	35,218
klasa 4	0,3375	0,2729	0,0848	0,3048	3,944	10,454	32,217
klasa 5	0,4973	0,1259	0,0641	0,3127	4,2017	17,1536	36,666

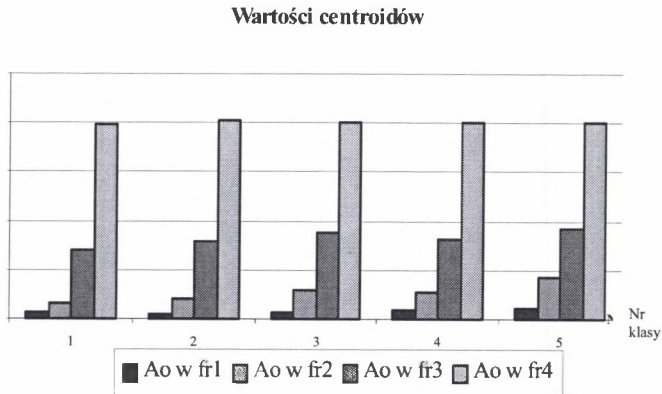
Wartości centroidów



Rys. 1. Udziały frakcji węgla dla centroidów

Fig. 1. Participation of coal fractions in the centroids

Wyznaczone wartości właściwości węgla w poszczególnych klasach zależą tylko i wyłącznie od wartości próbek węgla tworzących daną klasę. Można wobec tego uzyskane wartości średnie potraktować jako reprezentacyjne dla próbek węgla znajdujących się w bazie danych. Sposób ich wyznaczenia umożliwia wykorzystanie uzyskanych wartości węgla w bazie wiedzy będącej zbiorem testowym algorytmu samouczącego się.



Rys. 2. Procentowe zawartości popiołu we frakcjach dla centroidów
 Fig. 2. Proportional ash contents in fractions in the centroids

Znajomość podziału węgla na partie ziaren (frakcje) o różnym ciężarze właściwym, czyli o różnej zawartości popiołu umożliwia skonstruowanie krzywych wzbogacalności podających zależność między zawartością popiołu a wychodem frakcji. Z przebiegu krzywych można orientować się w możliwościach wzbogacania danego gatunku węgla.

4. Zastosowanie wyznaczonych rzeczywistych klas węgla

Wyznaczone klasy węgla umożliwiają zastosowanie rozmytego systemu decyzyjnego lub sieci neuronowych. Każda z tych technik wymaga wiedzy eksperta pozyskiwanej podczas uczenia się modelu na podstawie zbioru testowego [10].

Przyjęte cele technologiczne i ekonomiczne zakładu (np. żądana przepustowość zakładu, wydajne zużycie odczynników, wymagana jakość produktów) definiują kryteria oceny sterowania zakładem. Cele te są uwzględniane przy tworzeniu tablicy decyzyjnej w procesie uczenia się. Częstotliwość powtarzania etapu uczenia się systemu uzupełniającego istniejące modele zależy przede wszystkim od zmiany celów produkcyjnych.

Identyfikacja w czasie rzeczywistym polega na porównaniu bieżących wartości opisujących próbkę z kombinacjami punktów reprezentujących kolejne klastry. W zależności od określonej aktualnej klasy surowca oraz aktualnie występującej klasy surowca podejmowana jest decyzja według przyjętej tablicy decyzyjnej. Powoduje to wprowadzenie zmian do procesu w sposób ostry typu crisp (natychmiastowa zmiana parametrów procesu do

konkretnie ustalonych wartości, np. zwiększenie przepływu surowca z 225 l/h do 250 l/h) lub w sposób rozmyty (przejsie do aktualnego stanu wg wartości funkcji przynależności).

W prezentowanym przykładzie przedstawiono prognozy cen w pięciu wyznaczonych rzeczywistych klasach węgla. Prognozy postawione w oparciu o ceny węgla wzorcowego według formuły Blaschkego prezentuje tablica 2. Prognozy postawione w oparciu o ceny węgla wskaźnicowego według formuły Blaschkego prezentuje tablica 3.

Tablica 2

Prognozy postawione w oparciu o ceny węgla wzorcowego

	y_1	y_2	y_3	y_4
Klasa 1	7,536	0,7145	1,1450	0,818
Klasa 2	7,175	0,6985	1,1480	0,802
Klasa 3	7,246	0,7090	1,1475	0,814
Klasa 4	9,948	0,6952	1,1205	0,779
Klasa 5	9,602	0,6873	1,1240	0,773

gdzie y_1 – zawartość popiołu w trzech frakcjach
 y_2 – wychody w trzech frakcjach
 y_3 – cena względna (część ceny węgla wzorcowego)
 y_4 – cena za 1 t nadawy (część ceny węgla wzorcowego)

Tablica 3

Prognozy postawione w oparciu o ceny węgla wskaźnicowego

	y_1	y_2	y_3
Klasa 1	0,740	0,7145	1,0357
Klasa 2	0,730	0,6985	1,0451
Klasa 3	0,750	0,7090	1,0578
Klasa 4	0,703	0,6952	1,0110
Klasa 5	0,700	0,6873	1,0185

gdzie y_1 – cena względna (część ceny węgla wskaźnicowego)
 y_2 – waga produktu w 1t nadawy
 y_3 – cena za 1t produktu (część ceny węgla wskaźnicowego)

5. Rozmyty model lingwistyczny (model Mamdaniego)

Modele lingwistyczne, stanowiące podstawę modelowania jakościowego, opisują zachowanie systemu za pomocą reguł JEŻELI-TO o nieokreślonych predykatkach, w których zawarta jest baza wiedzy o systemie uwzględniająca rozmytą lub nieprecyzyjną wiedzę.

Modele lingwistyczne pozwalają na podejmowanie decyzji w oparciu o mechanizmy wnioskowania rozmytego, imitujące mechanizmy wnioskowania ludzkiego.

Model lingwistyczny składa się ze zbioru reguł o postaci [4, 6, 11]:

$$R^{(j)}: \text{JEŻELI } x_1 \text{ jest } A_1^j \text{ i... i } x_n \text{ jest } A_n^j \text{ TO } y_1 \text{ jest } B_1^j \text{ i ... i } y_m \text{ jest } B_m^j \quad (5.1)$$

gdzie:

- $R^{(j)}$ oznacza j -tą regułę w bazie reguł, $j = 1, \dots, J$,
- $x = (x_1, x_2, \dots, x_n)^T$ jest wektorem zmiennych wejściowych lingwistycznych zdefiniowanych w przestrzeni wejściowej, $x \in X \subset R^n$,
- A_1^j, \dots, A_n^j ($j = 1, 2, \dots, J$) są zbiorami rozmytymi (wartościami lingwistycznymi) zdefiniowanymi w przestrzeni zmiennych wejściowych,
- $y = (y_1, y_2, \dots, y_m)^T$ jest wektorem zmiennych wyjściowych, $y \in Y \subset R^m$,
- B_1^j, \dots, B_m^j są zbiorami rozmytymi (wartościami lingwistycznymi) zdefiniowanymi w przestrzeni zmiennych wyjściowych.

Zbiory rozmyte w przestrzeni zmiennych wejściowych i wyjściowych są scharakteryzowane przez odpowiednie funkcje przynależności:

$\mu_{A_i^j}(x_i) : R \rightarrow [0,1]$ jest funkcją przynależności zbioru rozmytego A_i^j .

$\mu_{B_k^j}(y_k) : R \rightarrow [0,1]$ jest funkcją przynależności zbioru rozmytego B_k^j .

W modelach Mamdaniego reguły rozmyte *JEŻELI-TO* interpretuje się jako rozmyte implikacje, zdefiniowane na produkcie kartezjańskim $X \times Y$ przestrzeni wejście-wyjście. Dla modelu MISO, o jednym wyjściu można to zapisać:

$$R^j: A_1^j \times \dots \times A_n^j \rightarrow B^j \quad (5.2)$$

z odpowiednimi funkcjami przynależności:

$$\mu_{A^j \rightarrow B^j}(x, y) = \min(\mu_{A^j}(x), \mu_{B^j}(y)) \quad (5.3)$$

$$\text{gdzie} \quad \mu_{A^j}(x) = \min(\mu_{A_1^j}(x_1), \dots, \mu_{A_n^j}(x_n)) \quad (5.4)$$

Agregacja reguł $R^{(j)}$ w całej przestrzeni $X \times Y$ jest scharakteryzowana przez spójnik *TAKŻE (also)*, z uwzględnieniem funkcji przynależności.

Algorytm podejmowania decyzji (sterowania, predykcji), dla określonego już modelu Mamdaniego i pewnej, zmierzonej wartości x_0 , jaką przyjmie wartość wejściowa, potraktowanej jako tzw. sigleton, można zapisać w postaci następującej:

dla każdej j -tej reguły wylicz poziom zapłonu zmiennej wejściowej:

$$z_j = \mu_{A_j}(x_0) \quad (5.5)$$

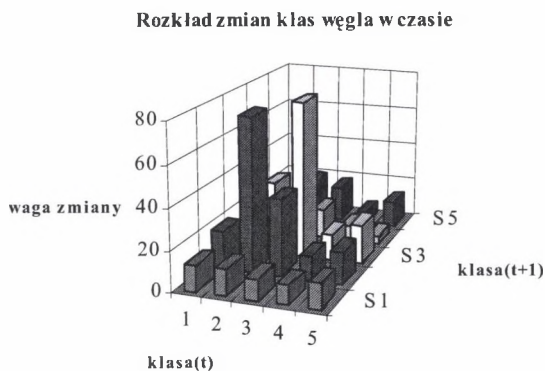
określ zbiór rozmyty B^j_0 zmiennej wyjściowej y , odpowiadający wartości wejściowej x_0 , przeprowadź agregację zbiorów rozmytych B^j_0 , stosując operację sumy (*max*).

W konstruowaniu modeli rozmytych w oparciu o zbiór obserwacji wejścia i wyjścia badanego systemu, często wykorzystuje się powyższe typy modeli [4, 6, 11]. Można stosować bardzo szeroką skalę modeli typu wejście-wyjście o dyskretnych wartościach czasu, np.:

$$R^{(i)}: w_j \text{ (JEŻELI } y(k) \text{ jest } A_1^j \text{ i } y(k-1) \text{ jest } A_2^j \text{ i } x(k) \text{ jest } A_3^j \text{ i ... i } x(k-n+3) \text{ jest } A_n^j \text{, TO } y(k+1) \text{ jest } B^j) \quad (5.6)$$

w których zmienne wejściowe i wyjściowe są zmiennymi rozmytymi (lingwistycznymi).

W sklasyfikowanym zbiorze danych opisującym właściwości węgla surowego wyznaczono 5 rzeczywistych klas. Celem jest stworzenie modelu opisanego funkcją $f(x(c=1), \dots, x(c=5))$. Zbiór wartości lingwistyczne w dziedzinie poprzednika $Klasa(n)$ oraz następnika $Klasa(n+1)$ jest zdefiniowany jako zbiór numerów wyznaczonych klas węgla. Rozmycie zmiennych lingwistycznych jest dokonane arbitralnie z wykorzystaniem trójkątnych, symetrycznych funkcji przynależności odpowiednio $\mu_a(x)$, $\mu_b(x)$, $\mu_c(x)$, w granicach dla zmiennej wejściowej (poprzednika) $X(n)$. Określono siatkę składającą się z 25 obszarów $A_1 \times S_1$, $A_2 \times S_1$, ..., $A_5 \times S_5$. Liczby par $\{Klasa(n), Klasa(n+1)\}$ w poszczególnych obszarach zostały przedstawione na rysunku 3 oraz w tabelicy 4.



Rys. 3. Rozkład zmian występowania klas węgla w czasie

Fig. 3. Distribution of changes of the coal classes in time

Tablica 4

Liczba zmian klas węgla w czasie

	Klasa 1	Klasa 2	Klasa 3	Klasa 4	Klasa 5
Klasa 1	13	13	10	9	12
Klasa 2	19	76	38	12	16
Klasa 3	14	36	78	13	19
Klasa 4	5	16	15	8	4
Klasa 5	7	20	18	6	13

W tablicy 5 przedstawiono prawdopodobieństwo warunkowe, że dla poprzednika A_i ma miejsce następnik S_j .

Tablica 5

Prawdopodobieństwo warunkowe zmian klas węgla w czasie

	Klasa 1	Klasa 2	Klasa 3	Klasa 4	Klasa 5
Klasa 1	0,0265	0,0286	0,0224	0,0184	0,0245
Klasa 2	0,0388	0,1551	0,0776	0,0245	0,0327
Klasa 3	0,0286	0,0735	0,1408	0,0265	0,0388
Klasa 4	0,0102	0,0327	0,0306	0,0163	0,0082
Klasa 5	0,0143	0,0408	0,0367	0,0122	0,0265

Według teorii Dempstera –Shafera [9], prawdopodobieństwo warunkowe, że dla poprzednika A_i ma miejsce następnik S_j określa wagę reguły w_{ij} na obszarze $A_i \times S_j$, przy czym, biorąc pod uwagę prawdopodobieństwo warunkowe, zachodzi równość:

$w_{i1} + w_{i2} + w_{i3} = 1$; $i = 1, 2, 3$. Wobec tego po znormalizowaniu wartości prawdopodobieństwa warunkowego uzyskano wartości wag reguł. Wartości wag reguł przedstawia tablica 6.

Tablica 6

Wartości wag reguł dla zmian klas węgla w czasie

	Klasa 1	Klasa 2	Klasa 3	Klasa 4	Klasa 5
Klasa 1	0,22	0,24	0,19	0,15	0,20
Klasa 2	0,12	0,47	0,24	0,07	0,10
Klasa 3	0,08	0,24	0,46	0,09	0,13
Klasa 4	0,11	0,33	0,31	0,17	0,08
Klasa 5	0,12	0,31	0,28	0,09	0,20

Baza reguł modelu rozmytego składa się z następujących formuł elementarnych:

JEŻELI X_i jest w 1. klasie TO X_{t+1} jest w 1. klasie z wagą $n_{11} = 0,22$

TAKŻE JEŻELI X_i jest w 1. klasie TO X_{t+1} jest w 2. klasie z wagą $n_{12} = 0,24$

TAKŻE JEŻELI X_i jest w 1. klasie TO X_{t+1} jest w 3. klasie z wagą $n_{13} = 0,19$

TAKŻE JEŻELI X_i jest w 1. klasie TO X_{t+1} jest w 4. klasie z wagą $n_{14} = 0,15$

TAKŻE JEŻELI X_i jest w 1. klasie TO X_{t+1} jest w 5. klasie z wagą $n_{15} = 0,20$

TAKŻE JEŻELI X_i jest w 2. klasie TO X_{t+1} jest w 1. klasie z wagą $n_{21} = 0,12$

TAKŻE	JEŻELI X_t	jest w 2. klasie	TO X_{t+1}	jest w 2. klasie z wagą $n_{22} = 0,47$
TAKŻE	JEŻELI X_t	jest w 2. klasie	TO X_{t+1}	jest w 3. klasie z wagą $n_{23} = 0,24$
TAKŻE	JEŻELI X_t	jest w 2. klasie	TO X_{t+1}	jest w 4. klasie z wagą $n_{24} = 0,07$
TAKŻE	JEŻELI X_t	jest w 2. klasie	TO X_{t+1}	jest w 5. klasie z wagą $n_{25} = 0,10$
TAKŻE	JEŻELI X_t	jest w 3. klasie	TO X_{t+1}	jest w 1. klasie z wagą $n_{31} = 0,08$
TAKŻE	JEŻELI X_t	jest w 3. klasie	TO X_{t+1}	jest w 2. klasie z wagą $n_{32} = 0,24$
TAKŻE	JEŻELI X_t	jest w 3. klasie	TO X_{t+1}	jest w 3. klasie z wagą $n_{33} = 0,46$
TAKŻE	JEŻELI X_t	jest w 3. klasie	TO X_{t+1}	jest w 4. klasie z wagą $n_{34} = 0,09$
TAKŻE	JEŻELI X_t	jest w 3. klasie	TO X_{t+1}	jest w 5. klasie z wagą $n_{35} = 0,13$
TAKŻE	JEŻELI X_t	jest w 4. klasie	TO X_{t+1}	jest w 1. klasie z wagą $n_{41} = 0,11$
TAKŻE	JEŻELI X_t	jest w 4. klasie	TO X_{t+1}	jest w 2. klasie z wagą $n_{42} = 0,33$
TAKŻE	JEŻELI X_t	jest w 4. klasie	TO X_{t+1}	jest w 3. klasie z wagą $n_{43} = 0,31$
TAKŻE	JEŻELI X_t	jest w 4. klasie	TO X_{t+1}	jest w 4. klasie z wagą $n_{44} = 0,17$
TAKŻE	JEŻELI X_t	jest w 4. klasie	TO X_{t+1}	jest w 5. klasie z wagą $n_{45} = 0,08$
TAKŻE	JEŻELI X_t	jest w 5. klasie	TO X_{t+1}	jest w 1. klasie z wagą $n_{51} = 0,12$
TAKŻE	JEŻELI X_t	jest w 5. klasie	TO X_{t+1}	jest w 2. klasie z wagą $n_{52} = 0,31$
TAKŻE	JEŻELI X_t	jest w 5. klasie	TO X_{t+1}	jest w 3. klasie z wagą $n_{53} = 0,28$
TAKŻE	JEŻELI X_t	jest w 5. klasie	TO X_{t+1}	jest w 4. klasie z wagą $n_{54} = 0,09$
TAKŻE	JEŻELI X_t	jest w 5. klasie	TO X_{t+1}	jest w 5. klasie z wagą $n_{55} = 0,20$

6. Podsumowanie

Rzeczywiste klasy węgla wyznaczone zostały w oparciu o wyniki pomiarów właściwości jakościowych węgla. Wobec tego wiarygodnie charakteryzują zbiór próbek węgla i mogą posłużyć jako zbiór testowy dla algorytmów samouczących się.

Zastosowany algorytm grupowania rozmytego posłużył wykonaniu wielowymiarowej analizy właściwości próbek węgla wyznaczając rzeczywiste wzorce danych. Utworzono żadaną liczbę klas poprzez wyliczenie współczynników ich centroidów. Na podstawie współczynników centroidów wyznaczono prognozy możliwych do uzyskania cen węgla w oparciu o węgiel wskaźnikowy lub węgiel wzorcowy.

Rozkład wiarygodności klas węgla określa wagi reguł modelu rozmytego typu Mamdaniego. Na podstawie współczynników centroidów i stosując formułę Blaschkego można wyznaczyć wskaźniki cenowe węgla dla kolejnych klas.

LITERATURA

1. Bezdek J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. New York 1981.

2. Blaschke W.: System cen energetycznych węgla kamiennego. Studia Rozprawy Monografic 77. Wydawnictwo Instytutu Gospodarki Surowcami Mineralnymi i Energią PAN, Kraków 2000.
3. Blaschke W., Mokrzycki E., Zhongjian Shan: Coal preparation economics, New Trends in Coal Preparation Technologies and Equipment. Gordon and Breach Publishers, 1994.
4. Czogała E., Frączek J., Sułkowski J.: Simulation of fuzzy control of the ventilation process in the longwalls with methan emission (Polish). *Archiwum Informatyki Teoretycznej i Stosowanej*, T.7 (1995).
5. Everitt B. S.: Cluster Analysis. Americas. Halsted Press, New York 1993.
6. Hellendoorn H, Driankov D.(Eds.): Fuzzy model identification. Springer - Verlag, Heidelberg 1997.
7. Klir G. J., Yuan Bo.: Fuzzy Sets and Fuzzy Logic. Prentice Hall PTR. Upper Saddle River. New Jersey 1995.
8. Kowal A., Walaszek-Babiszewska A.: Koncepcja budowy modeli rozmytych w oparciu o dane eksperymentalne. *Górnictwo Zrównoważonego Rozwoju. Konferencja III Energooszczędne i Niezawodne Maszyny Górnicze. Zeszyty Naukowe Górnictwo*, 246, Gliwice 2000.
9. Walaszek-Babiszewska A.: Stochastic and fuzzy modeling for sampling research. *Akustyka Molekularna i Kwantowa (Molecular & Quantum Acoustics)* 1997.
10. Zadeh L.: From computing with numbers to computing with words – from manipulation of measurements to manipulation of perceptions. *IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications* 1999.

Recenzent: Dr hab. inż. Kazimierz Jaracz, prof. nzw. w AP w Krakowie

Abstract

In the article an example of the real coal clusters delimitation is presented. The algorithm of fuzzy date classification is used to make partitions of date.

In the introduction the concept of the coal quality is presented. In the second part the basic features of fuzzy crisp clustering methods are mentioned. In the third part of the article the algorithm of fuzzy date classification is used to make partitions of date. A file of 500 coal samples is divided. Results of the partitions are presented as the participation of coal fractions in the centroids (fig.1) and as proportional ash contents in fractions in the centroids (fig.2).

In the next parts, the prognostic prices based on the clusters is presented. The Blaschke formula is used to delimitation of the indicatory and standard prices. The prognostic is necessary to create a fuzzy linguistic Mamdani model. The rule base is the result of this modelling.