

Iwona Staniec, Jan Żółtowski

Politechnika Łódzka

DANE SYMBOLICZNE W KLASYFIKACJI OCZEKIWAŃ PRACODAWCÓW WOBEC ABSOLWENTÓW KIERUNKÓW INŻYNIERSKICH I MENEDŻERSKICH*

1. Wstęp

Znajomość oczekiwań pracodawców powinna ułatwić absolwentom start zawodowy. Dlatego coraz więcej czasu poświęca się przygotowaniu profili kształcenia tak, aby odpowiadały one oczekiwaniom pracodawców. Właściwe przygotowanie absolwenta jest zjawiskiem nietrywialnym ze względu na zmieniające się potrzeby rynku pracy oraz wymagania pracodawców. Wymaga ono również porównywalnych kwalifikacji, a porównywalność jest możliwa tylko po wprowadzeniu powszechnie akceptowanych norm.

Charakterystyki oczekiwań pracodawców mogą być przedstawione jako zmienne symboliczne w postaci wariantów cech z przyporządkowanymi im wagami. Celem przedstawionych rozważań jest klasyfikacja kierunków studiów ze względu na ocenę umiejętności absolwentów dokonaną przez pracodawców. Dzięki niej możliwe jest wskazanie cech charakteryzujących poszczególne grupy oraz wyjaśnienie, jakimi czynnikami różnią się wyodrębnione grupy kierunków. Podstawą interpretacji będą zmienne biorące udział w procesie klasyfikacji. W przedstawionych rozważaniach w klasyfikacji absolwentów wykorzystano specyficzne dane.

2. Zmienne symboliczne

W rozważanym przypadku wykorzystano zmienną symboliczną¹, która jest odwzorowaniem przyporządkowującym poszczególnym obiektom kolejne warianty zmiennej z ustalonymi wagami, czyli:

¹ Dyskusyjny pozostaje problem nazewnictwa tak określonych zmiennych, por. prace [Wilk 2006, s. 224-236; Machowska-Szewczyk 2008, s. 240-248; Malerba, Esposito, Monopoli 2002, s. 31-33; Walesiak 1983, s. 26-28].

$$X_i : O_k \rightarrow \left((x_{i1}, p_k(x_{i1})), \dots, (x_{im}, p_k(x_{im})) \right),$$

gdzie: $i \in \{1, 2, \dots, K\}$ i $\sum_{j=1}^m p_k(x_{ij}) = 1$.

Zmienne z wagami przez swój specyficzny charakter i nietypową strukturę uniemożliwiają zastosowanie klasycznych miar odległości, a także miar przeznaczonych do analizy podobieństwa obiektów opisanych innymi typami zmiennych. W tym przypadku pojedynczemu obiektowi jest przypisany więcej niż jeden wariant zmiennej, a wariantom przypisane są wagi $p_k(x_{ij})$ – częstości występowania wariantu x_{ij} [Brito, de Carvalho 2002, s. 13].

Podobieństwo obiektów O_k oraz O_l względem zmiennej z wagami mierzy się przez porównanie rozkładów prawdopodobieństwa zmiennej dla poszczególnych obiektów [Malerba, Esposito, Monopoli 2002, s. 33]. Dwa obiekty są więc tym bardziej podobne, im mniej różnią się co do wartości wag poszczególnych wariantów, tzn. im bardziej proporcjonalne są ich rozkłady prawdopodobieństwa. Do oceny podobieństwa między obiektami O_k oraz O_l względem zmiennej X_i można wykorzystać m.in. odległość Minkowskiego:

$$d_p^i(O_k, O_l) = \left(\sum_{j=1}^m |p_k(x_{ij}) - p_l(x_{ij})|^p \right)^{\frac{1}{p}}, \quad (1)$$

gdzie: $p \geq 1$.

Kolejnym etapem jest wyznaczenie funkcji odległości dla zdefiniowanych obiektów strukturalnych. Stosowane w praktyce miary odległości zmiennych są w większości adaptacją miar sformułowanych dla rozkładów prawdopodobieństwa oraz klasycznych miar podobieństwa obiektów, a także miar wykorzystywanych w segmentacji obrazów (por. [Wilk 2006, s. 228]). W przedstawionych rozważaniach odległość między dwoma tak zdefiniowanymi obiektami jest zagregowaną wartością odległości Minkowskiego obiektów względem kolejnych zmiennych opisaną poniższym wzorem [Malerba, Esposito, Monopoli 2002, s. 36]:

$$d_p^i(O_k, O_l) = \left(\sum_{i=1}^K (\omega_i \cdot d^i(O_k, O_l))^p \right)^{\frac{1}{p}}, \quad (2)$$

gdzie: $\omega_i \in [0; 1]$ jest wagą odległości obiektów względem zmiennej X_i , a $\sum_{i=1}^K \omega_i = 1$.

Po wyznaczeniu macierzy odległości możliwe jest zastosowanie klasycznych metod klasyfikacji. Jedną z najpopularniejszych jest metoda Warda [Ward 1963], w której procedura grupowania składa się z $n-1$ kolejnych kroków, gdzie n stanowi liczbę elementów wyjściowego zbioru. Na etapie zerowym tworzone są grupy jed-

noelementowe ze wszystkich elementów badanej zbiorowości. W kolejnych krokach algorytmu łączone są dwie grupy (z powstałych w poprzednim kroku procedury) w jedną grupę tak, aby zminimalizować sumę kwadratów odchyleń wszystkich obiektów z tych dwóch grup od środka ciężkości nowej grupy, która powstanie w wyniku połączeń tych dwóch grup (na każdym etapie łączenia grup obiektów ze wszystkich możliwych do łączenia grup obiektów łączy się w jedną grupę te grupy, które w rezultacie tworzą grupę obiektów o najmniejszym zróżnicowaniu ze względu na opisujące je zmienne). W wyniku zastosowania powyższej procedury zgrupowane zostają wszystkie elementy w jedną całość. W sytuacjach, gdy chce się zakończyć procedurę przed uzyskaniem jednej grupy, za kryterium zatrzymania przyjmuje się odpowiedni (czasami subiektywnie wyznaczony) wskaźnik poziomu odcięcia. Problem doboru poziomu odcięcia można ominąć, stosując indeks spójności grup, którego wybór jest kwestią sporną. Dyskusję na ten temat przeprowadzono w pracach [Halkidi i in. 2001; Migdał-Najman, Najman 2005]. Najpopularniejszym ze stosowanych obecnie jest indeks Calińskiego-Harabasz [1974]:

$$CH(N_c) = \frac{tr(B) \cdot (n - N_c)}{(N_c - 1) \cdot tr(W)}, \quad (3)$$

gdzie: $tr(B)$ – ślad macierzy międzyklasowej,
 $tr(W)$ – ślad macierzy wewnątrzklasowej,
 N_c – liczba klas w danym kroku analizy,
 n – liczba danych poddanych analizie.

Według tego indeksu na każdym kroku analizy wyznacza się ślad macierzy kowariancji międzyklasowej oraz ślad macierzy kowariancji wewnątrzklasowej. Za optymalną liczbę klas uznaje się tę, dla której wyznaczony indeks CH osiąga wartość max (największą).

3. Analizy numeryczne

W drugiej połowie 2007 r. w Katedrze Zarządzania Politechniki Łódzkiej przeprowadzono badania ankietowe dotyczące zapotrzebowania na pracę oraz kwalifikacji inżynierów i menedżerów w województwie łódzkim. Badania ankietowe przeprowadzono na próbie losowo wybranych 500 pracodawców z województwa łódzkiego – każdy z nich oceniał umiejętności zatrudnianych absolwentów. W analizie wykorzystano dane uzyskane w wyniku tego badania, które dotyczyły oceny przygotowania absolwentów:

- w zakresie wiedzy kierunkowej: znajomość specjalistycznej wiedzy (pojęć procesów) – x_1 ; umiejętność rozwiązywania problemów typowych dla dziedziny – x_2 ; obsługi specjalistycznej aparatury, maszyn i urządzeń – x_3 ; obsługi specjalistycznych programów komputerowych – x_4 ; prowadzenia projektów, pracy w zespole projektowym – x_5 ; znajomość regulacji prawnych dotyczących dziedziny, np. norm

jakości UE z dziedziny – x_6 ; analizy techniczno-ekonomicznej (np. kosztorysowanie, harmonogramowanie) – x_7 ; działalności badawczo-rozwojowej – x_8 ;

- w zakresie wiedzy ogólnej: języki obce – x_9 ; ekonomika, organizacja przedsiębiorstw, finanse, marketing, handel – x_{10} ; zarządzanie jakością, produkcją, logistyka – x_{11} ; materiałoznawstwo – x_{12} ; ochrona środowiska – x_{13} ; ergonomia – x_{14} ; potrzeby niepełnosprawnych – x_{15} ; etyka – x_{16} ; obsługa komputera – x_{17} ;
- w zakresie umiejętności miękkich: porozumiewanie się ze specjalistami z własnej dziedziny – x_{18} ; porozumiewanie się ze specjalistami z innych dziedzin – x_{19} ; porozumiewanie się z niespecjalistami (personelem) – x_{20} ; prowadzenie negocjacji – x_{21} ; praca w grupie – x_{22} ; motywowanie, kierowanie – x_{23} ; samodzielność – x_{24} ; kreatywność – x_{25} ; analityczne myślenie – x_{26} ; poszukiwanie i gromadzenie informacji – x_{27} ; świadomość samokształcenia – x_{28} .

Ocenę stopnia przygotowania absolwentów w trzystopniowej skali (dobrze, średnio, źle) pracodawcy przeprowadzili dla następujących kierunków studiów (szerzej na ten temat w pracach [Staniec, Żółtowski 2008; Staniec 2008; Szmit, Staniec, Żółtowski 2008]):

- inżynieria chemiczna i inżynieria środowiska,
- automatyka i robotyka, elektrotechnika, elektronika i telekomunikacja,
- technologia chemiczna, chemia,
- technologia żywności, biotechnologia, ochrona środowiska, inżynieria biochemiczna,
- wszystkie specjalności menedżerskie,
- budownictwo,
- włókiennictwo,
- informatyka,
- mechanika.

Ocena stopnia przygotowania absolwentów jest mierzona na słabej skali. Użyta skala pomiarowa ma wpływ na liczbę informacji dostarczanych przez zmienną. W opisanym przypadku ocenę umiejętności absolwenta dokonaną przez pracodawców przedstawiono jako specyficzny rodzaj zmiennej symbolicznej, gdzie każda z nich przyjmuje różne warianty, a poszczególnym wariantom przypisane są wagi w postaci częstości występowania. Każdy z wariantów został przedstawiony w postaci listy kategorii porządkowych. Dla każdego obiektu określono zestawy zmiennych i odpowiadające im wagi. Przykładowe zmienne symboliczne dla absolwentów kierunku technologia żywności, biotechnologia, ochrona środowiska, inżynieria biochemiczna przedstawiono w tab. 1.

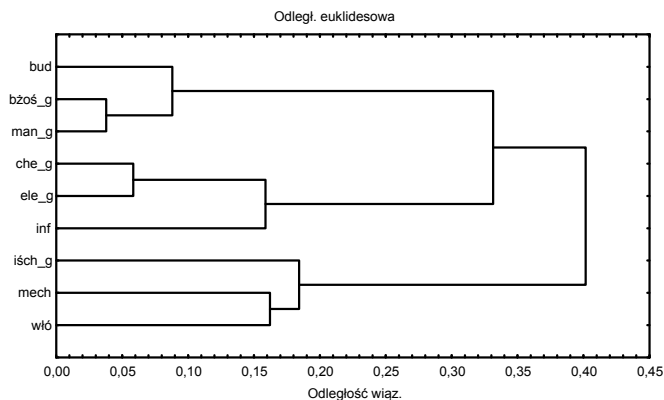
Następnie wyznaczono dla analizowanych kierunków studiów miary podobieństwa opisane regułą (1) i za ich pomocą wyznaczono macierz odległości, której elementy są zagregowaną wartością odległości Minkowskiego analizowanych kierunków studiów względem kolejnych zmiennych opisaną wzorem (2). Dla macierzy odległości Minkowskiego zastosowano do klasyfikacji umiejętności absolwentów poszczególnych kierunków metodę Warda. Powstały dendrogram pozwala określić odległości między skupieniami oraz wyodrębnić zarówno elementy najbardziej

Tabela 1. Zmienne symboliczne dla absolwentów kierunku technologia żywności, biotechnologia, ochrona środowiska, inżynieria biochemiczna

Nr	Zmienna	Warianty zmiennej z udziałem procentowym					
x ₁	znajomość specjalistycznej wiedzy (pojęć procesów)	dobrze	66,67%	średnio	33,33%	źle	0,00%
x ₂	rozwiązywanie problemów typowych dla dziedziny	dobrze	66,67%	średnio	25,00%	źle	8,33%
x ₃	obsługa specjalistycznej aparatury, maszyn i urządzeń	dobrze	66,67%	średnio	33,33%	źle	0,00%
x ₄	obsługa specjalistycznych programów komputerowych	dobrze	84,62%	średnio	7,69%	źle	7,69%
x ₅	przewodzenie projektów, praca w zespole projektowym	dobrze	72,73%	średnio	27,27%	źle	0,00%
x ₆	znajomość regulacji prawnych dotyczących dziedziny, np. norm jakości UE z dziedziny	dobrze	40,00%	średnio	60,00%	źle	0,00%
x ₇	analizy techniczno-ekonomiczne (np. kosztorysowanie, harmonogramowanie)	dobrze	62,50%	średnio	37,50%	źle	0,00%
x ₈	działalność badawczo-rozwojowa	dobrze	50,00%	średnio	50,00%	źle	0,00%
x ₉	języki obce	dobrze	53,85%	średnio	38,46%	źle	7,69%
x ₁₀	ekonomika, organizacja przedsiębiorstw, finanse, marketing, handel	dobrze	40,00%	średnio	60,00%	źle	0,00%
x ₁₁	zarządzanie jakością, produkcją, logistyka	dobrze	33,33%	średnio	66,67%	źle	0,00%
x ₁₂	materiałoznawstwo	dobrze	71,43%	średnio	28,57%	źle	0,00%
x ₁₃	ochrona środowiska	dobrze	100,00%	średnio	0,00%	źle	0,00%
x ₁₄	ergonomia	dobrze	50,00%	średnio	50,00%	źle	0,00%
x ₁₅	potrzeby niepełnosprawnych	dobrze	28,57%	średnio	71,43%	źle	0,00%
x ₁₆	etyka	dobrze	50,00%	średnio	50,00%	źle	0,00%
x ₁₇	obsługa komputera	dobrze	90,91%	średnio	9,09%	źle	0,00%
x ₁₈	porozumiewanie się ze specjalistami z własnej dziedziny	dobrze	92,31%	średnio	7,69%	źle	0,00%
x ₁₉	porozumiewanie się ze specjalistami z innych dziedzin	dobrze	84,62%	średnio	15,38%	źle	0,00%
x ₂₀	porozumiewanie się z niespecjalistami (personelem)	dobrze	92,31%	średnio	7,69%	źle	0,00%
x ₂₁	przewodzenie negocjacji	dobrze	54,55%	średnio	45,45%	źle	0,00%
x ₂₂	praca w grupie	dobrze	76,92%	średnio	23,08%	źle	0,00%
x ₂₃	motywowanie, kierowanie	dobrze	75,00%	średnio	25,00%	źle	0,00%
x ₂₄	samodzielność	dobrze	84,62%	średnio	15,38%	źle	0,00%
x ₂₅	kreatywność	dobrze	69,23%	średnio	30,77%	źle	0,00%
x ₂₆	analityczne myślenie	dobrze	84,62%	średnio	15,38%	źle	0,00%
x ₂₇	poszukiwanie i gromadzenie informacji	dobrze	91,67%	średnio	8,33%	źle	0,00%
x ₂₈	świadomość samokształcenia	dobrze	100,00%	średnio	0,00%	źle	0,00%

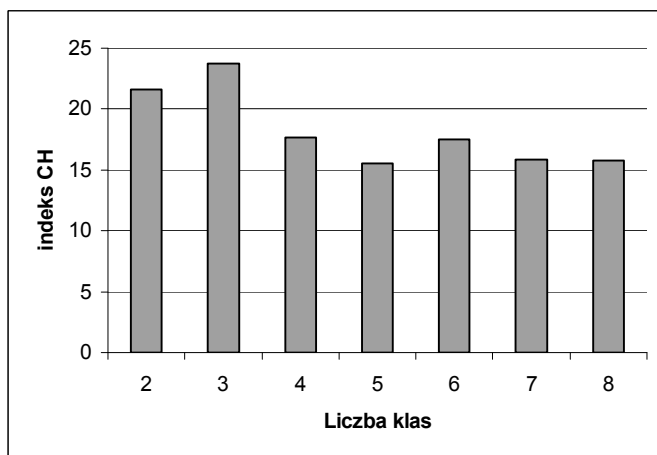
Źródło: opracowanie własne.

podobne w danej grupie, jak i elementy najmniej pasujące do skupienia pod względem zastosowanego kryterium grupowania. Analizując diagram (rys. 1), zauważono, że nie ma kierunków istotnie odróżniających się od pozostałych. Zatem dla wszystkich wyznaczono indeks Calińskiego-Harabasa według formuły (3), a jego wartości przedstawiono na rys. 2. Maksymalną wartość przyjmuje on dla trzech klas.



Rys. 1. Skupienia kierunków według Warda

Źródło: opracowanie własne.



Rys. 2. Wartości indeksu CH

Źródło: opracowanie własne.

Przeprowadzone analizy wskazują jednoznacznie na konieczność utworzenia trzech klas. Nazwy utworzonych klas określono na podstawie ocen umiejętności absolwentów poszczególnych kierunków studiów w trzech wyróżnionych grupach:

wiedza kierunkowa, ogólna i umiejętności miękkie, dokonanych przez pracodawców. W wyniku klasyfikacji otrzymano trzy klasy:

- **klasa 1: dobrzy absolwenci**
 - technologia żywności, biotechnologia, ochrona środowiska, inżynieria biochemiczna (boś_g),
 - wszystkie specjalności menedżerskie (man_g),
 - budownictwo (bud);
- **klasa 2: przeciętni absolwenci**
 - technologia chemiczna, chemia (che_g),
 - informatyka (inf),
 - automatyka i robotyka, elektrotechnika, elektronika i telekomunikacja (ele_g);
- **klasa 3: słabi absolwenci**
 - inżynieria chemiczna i inżynieria środowiska (iśch_g),
 - włókiennictwo (włó),
 - mechanika (mech).

Na podstawie otrzymanych klas można wyciągnąć następujące wnioski:

- klasa 1: dobrzy absolwenci – to absolwenci oceniani dobrze przez pracodawców pod względem wiedzy kierunkowej, wiedzy ogólnej oraz umiejętności miękkich,
- klasa 2: przeciętni absolwenci – to absolwenci oceniani dobrze przez pracodawców pod względem wiedzy kierunkowej, średnio pod względem wiedzy ogólnej oraz źle pod względem umiejętności miękkich,
- klasa 3: słabi absolwenci – to absolwenci oceniani średnio przez pracodawców pod względem wiedzy kierunkowej, źle pod względem wiedzy ogólnej i umiejętności miękkich. Warto zaznaczyć, iż na kierunki z tej grupy wybierają się najslabsi absolwenci szkół średnich, nie cieszą się one również popularnością, co widać już w procesie rekrutacji.

Porównując otrzymane wyniki z programami studiów realizowanymi na poszczególnych kierunkach, stwierdzono, że:

- w programach kształcenia na kierunkach, które zgrupowano w klasie 1, znajdują się nie tylko przedmioty kierunkowe, ale również przedmioty pozwalające na rozszerzenie wiedzy ogólnej i te, które rozwijają i oceniają umiejętności miękkie,
- w programach kształcenia na kierunkach, które zgrupowano w klasie 2, znajdują się przede wszystkim przedmioty kierunkowe i wymagane przez standardy kształcenia przedmioty ogólne, brak zaś w nich przedmiotów rozwijających umiejętności miękkie – na które pracodawcy coraz częściej zwracają uwagę,
- w programach kształcenia na kierunkach, które zgrupowano w klasie 3, znajdują się przedmioty kierunkowe i specjalnościowe oraz wymagane przez standardy kształcenia przedmioty ogólne, brak w nich przedmiotów pozwalających na rozwijanie umiejętności miękkich. Percepcja studentów na tych kierunkach nie pozwala na przyswojenie oferowanej wiedzy w pełni. Średnia na dyplomach na Politechnice Łódzkiej dla tych kierunków jest najniższa.

4. Podsumowanie

Wyniki badań wskazują, iż pracodawcy oceniają absolwentów dobrze pod względem konkretnych umiejętności związanych z danym kierunkiem, zatem programy kształcenia przygotowują dobrze w zakresie wiedzy kierunkowej. Jednak obecnie pracodawcy oczekują od absolwentów również wiedzy wykraczającej poza standardy kształcenia. Wynika to przede wszystkim z funkcjonowania w dynamicznie zmieniającym się otoczeniu, rozwoju nauki, techniki i technologii itp. Zatem konieczne jest wprowadzenie do programów kształcenia umiejętności miękkich, tj. umiejętności interpersonalnych i predyspozycji osobowościowych niezbędnych (lub przydatnych) w pracy zawodowej, np.: komunikatywności, umiejętności pracy w grupie, systematyczności, kreatywności.

Wydaje się również, że wykorzystanie ocen jako zmiennych z wagami w postaci częstości wystąpień pozwoliło na otrzymanie jednorodnych klas, wyraźne określenie cech wspólnych i różnicujących dla tych klas oraz odniesienie się również do programów kształcenia.

Literatura

- Brito P., De Carvalho F. (2002), *Symbolic clustering of constrained probabilistic data*, [w:] *Exploratory data analysis in empirical research*, red. O. Opitz, M. Schvaiger, Series „Studies in Classification, Data Analysis and Knowledge Organization”, Springer Verlag, Heidelberg, s. 12-21.
- Calinski T., Harabasz J. (1974), *A dendrite method for cluster analysis*, „Communication in Statistics” vol. 3, s. 1-27.
- Gatnar E. (1998), *Strukturalne metody klasyfikacji danych*, PWN, Warszawa.
- Halkidi M., Batistakis Y., Vazirgiannis M. (2001), *On clustering validation techniques*, „Journal of Intelligent Information Systems”, 17, no 2/3, s. 107-145.
- Machowska-Szewczyk M. (2008), *Porównanie wyników klasyfikacji nauczycieli akademickich pod względem wpływu osobowości na popularność wykładów ze statystyki*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 7 (1207), Taksonomia 15, AE, Wrocław, s. 240-248.
- Malerba D., Esposito F., Monopoli M. (2002), *Comparing dissimilarity measures for probabilistic symbolic objects*, [w:] *Data Mining III*, red. A. Zanasi, C.A. Brebbia, N.F.F. Ebecken, P. Melli, Series Management Information Systems, vol. 6, 31-40, WIT Press, Southampton, UK.
- Migdał-Najman K., Najman K. (2005), *Analizy metod ustalania liczby skupień*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 6 (1076), Taksonomia 12, AE, Wrocław, s. 265-273.
- Staniec I. (2008), *Oczekiwania pracodawców w zakresie kompetencji absolwentów*, [w:] *Teoria i praktyka zarządzania rozwojem organizacji*, red. S. Lachiewicz, A. Zakrzewska-Bielawska, Monografie Politechniki Łódzkiej, Łódź, s. 323-330.
- Staniec I., Żółtowski J. (2008), *Ocena kwalifikacji zawodowych absolwentów*, [w:] *Wyzwania dla zarządzania zasobami ludzkimi w sytuacji przemian na rynku pracy*, red. A. Stankiewicz-Mróż, J. Lenzion, Politechnika Łódzka Media Press, Łódź, s. 27-36.
- Szmit A., Staniec I., Żółtowski J. (2008), *Analiza kwalifikacji kadry inżynierskiej i menedżerskiej w regionie łódzkim na podstawie badania ankietowego*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 7 (1207), Taksonomia 15, AE, Wrocław, s. 144-151.
- Walesiak M. (1983), *Podobieństwo wielkości (skali) oraz kształtu (formy) w złożonych badaniach strukturalnych*, „Wiadomości Statystyczne” nr 3, s. 26-28.

Ward J.H. (1963), *Hierarchical grouping to optimize an objective function*, „Journal of the American Statistical Association” no 58, 236.

Wilk J. (2006), *Miary odległości obiektów opisanych zmiennymi symbolicznymi z wagami*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 7 (1126), Taksonomia 13, AE, Wrocław, s. 224-236.

THE SYMBOLIC DATA IN THE CLASSIFICATION OF EXPECTATIONS OF EMPLOYERS FOR GRADUATES IN ENGINEERING AND MANAGEMENT

Summary

The aim of the deliberations is to classify graduates of various courses due to the assessment of their skills by employers, and pointing out common characteristics and features that diverse different courses. An original approach in the article refers to using rates as a specific type of variables. The results of the classification for the data presented allow to obtain homogeneous groups of courses and to clearly define common and differing features.