

Leonard Rozenberg

BADANIE ZDOLNOŚCI DYSKRYMINACYJNEJ WYBRANYCH WSKAŹNIKÓW FINANSOWYCH PRZEDSIĘBIORSTWA Z WYKORZYSTANIEM PODEJŚCIA ROZMYTEGO

STUDY ON DISCRIMINANT CAPABILITY FOR SELECTED FINANCIAL RATIOS OF THE COMPANY USING FUZZY APPROACH

Katedra Inżynierii Zarządzania, Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
ul. Żołnierska 49, 71-210 Szczecin, e-mail: leonard.rozenberg@zut.edu.pl

Summary. The paper is presented an analysis of selected financial indicators using fuzzy approach in terms of the use of the selected membership function in the enterprise information system. Also discussed was the matter of choosing the type of membership function, the selected its trapezoidal shape. Based on the graphical presentation of the membership function for the selected financial ratios conducted a discussion on their usefulness to the system of discrimination states the company. The study was conducted on financial data warehouses 55 Polish companies (110 records). The presented method does not depend on the dimensionality of the data warehouse or the complexity of the problem and is so flexible that it can be used not only in management sciences, or computer science, but also in the biological sciences, medical or other, depending on where you are not given in acute form.

Słowa kluczowe: analiza wskaźnikowa, baza danych, podejście rozmyte, system wspomaganie decyzji.

Key words: database, decision support system, fuzzy approach, ratio analysis.

WSTĘP

Istnieje wiele złożonych problemów, których opis (lub szerzej rozwiązanie) wymaga przetwarzania ogromnych ilości informacji. Obróbka i przechowywanie tak dużych objętościowo zbiorów danych stały się możliwe dzięki technologii baz danych, ale wymiarowość tych baz ogranicza kompleksową analizę, jak też wydobycie ukrytej w nich wiedzy, którą dalej wykorzystywać można np. w systemach ekspertowych, procedurach optymalizacyjnych, które pośrednio lub bezpośrednio dotyczą wspomaganie decyzji (Śmiałkowska 2009).

Celem tego artykułu jest zaprezentowanie prostej metody służącej do oceny potencjału dyskryminacyjnego wskaźników analizy finansowej przedsiębiorstwa. W prezentowanym artykule pokazano zastosowanie podejścia rozmytego, które wspiera proces ekstrakcji wiedzy z bazy danych. Obiektem badań były sprawozdania finansowe firm, czyli bilanse, rachunki zysków i strat oraz rachunki przepływów pieniężnych¹. Problem badawczy polegał na znalezieniu metody,

¹ Sprawozdania z przepływów pieniężnych były generowane na podstawie bilansów i rachunków zysków i strat dla dwóch kolejnych okresów sprawozdawczych.

która pozwoli na wygodne i proste analizowanie danych z tych sprawozdań, i to niezależnie od ich wymiarowości. Zaprezentowana metoda jest przystępna i powinna wspierać np. procesy wnioskowania.

W praktyce zarządzania i podejmowania decyzji na ogół istnieje potrzeba śledzenia i analizowania ogromnej ilości danych i informacji, najistotniejszych z punktu widzenia przedsiębiorstwa wskaźników, często zmiennych w czasie, nie tylko ilościowo, ale również jakościowo. Do tych wskaźników należą np.: wskaźnik bieżącej płynności, rentowność sprzedaży, rentowność kapitału całkowitego, cykl odnawiania zapasów, rentowność kapitału własnego, ale także trudne do liczbowej reprezentacji ryzyko, poziom zadowolenia klientów, jakość produkcji, jakość obsługi, wskaźnik zmienności zachowań klienta itp. (Rozenberg 2001).

W związku z postępującą globalizacją i wirtualizacją gospodarki, w przyszłości pojawić się mogą inne wskaźniki, które dziś wydają się nieistotne, ale zawsze proces podejmowania decyzji wymagać będzie dostępu do ogromnych ilości informacji i danych, w odpowiednim czasie i w formie użytecznej dla rozwiązywania rzeczywistych problemów związanych z organizacją przedsiębiorstwa.

Decydenci podejmują decyzje na podstawie specjalnych metod i modeli, poczynając od prostych modeli rachunkowych, które mogą być aplikowane przy użyciu prostych arkuszy kalkulacyjnych, na modelach wyboru optymalnego, czy paretooptymalnego lub też suboptymalnego wyboru kończąc. Wśród takich systemów można wyróżnić:

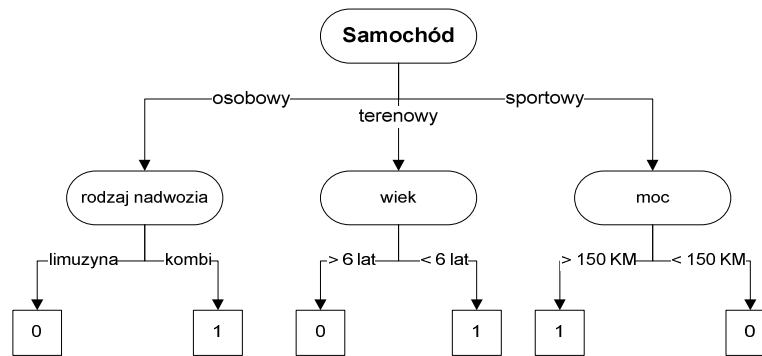
- tzw. systemy transakcyjne (TPS),
- informacyjne (*Management Information System* – MIS),
- wspomaganie podejmowania decyzji (*Decision Support Systems* – DSS),
- systemy informowania kierownictwa (*Execution Information Systems* – EIS).

Systemy wspomaganie decyzji, jak i systemy informowania kierownictwa EIS, to najczęściej systemy informatyczne zaprojektowane do uzyskania informacji niezbędnych w procesach podejmowania decyzji gospodarczych (biznesowych), które czerpać muszą wiedzę z niezliczonych źródeł.

Koncepcja wykorzystania do wspomaganie decyzji systemów informatycznych nie jest pomysłem ostatnich lat, ale dopiero od niedawna stała się możliwa ich efektywna implementacja. Wśród takich systemów, które w swojej istocie są przecież systemami reprezentującymi i odkrywającymi wiedzę, można wymienić:

1. Struktury drzewiaste, których koncepcja wykorzystania do reprezentacji wiedzy pojawiła się znacznie wcześniej niż samo pojęcie Data Mining'u czy nawet uczenia maszynowego. Opiera się ona na sekwencyjnym przechodzeniu przez „sito” warunków, aby na końcu otrzymać rozłączne zbiory spełniające określone, założone cechy. Formalnie drzewo decyzyjne jest acyklicznym, spójnym, skierowanym grafem. Krawędzie takiego grafu nazywamy gałęziami, wierzchołki natomiast, w zależności czy wychodzą z nich krawędzie, czy też nie – węzłami lub liśćmi.

Proste drzewo klasyfikujące według kryterium dychotomicznego, przyznające samochodom odpowiednie etykiety (**drogi** [1], **tani** [0]), przedstawiono na rysunku 2.



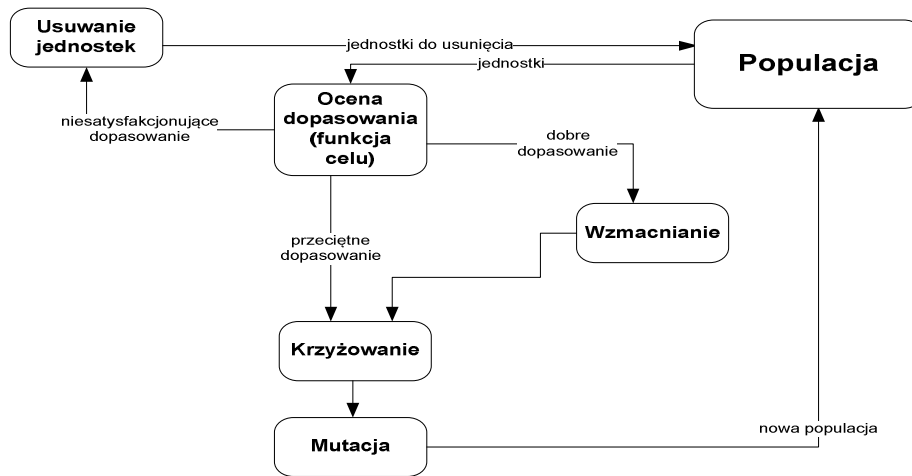
Rys. 1. Drzewo decyzyjne klasyfikujące samochody

Źródło: opracowanie własne.

2. Analiza statystyczna, która jest najstarszą i zarazem najlepiej poznaną metodą odkrywania wiedzy. Od początku jest ona dziedziną matematyki, więc trudno wskazać jednoznacznie moment jej powstania. Statystyka jest dzisiaj, patrząc na liczbę publikacji, niemodną dziedziną badań, ale w praktyce wszystkie stosowane obecnie metody ekstrakcji wiedzy w mniejszym lub większym stopniu bazują na jej osiągnięciach. Jeśli osiągnięcia statystyki nie są wykorzystywane bezpośrednio, jak np. w konstrukcji sztucznych sieci neuronowych, to wykorzystuje się je do doboru przykładów trenujących, jak i w ocenie trafności działania analizowanego algorytmu².
3. Algorytmy grupowania, czyli ogromny zbiór różnych procedur postępowania. Ich celem jest zamiana zbioru wszystkich przypadków w kilka, kilkanaście podzbiorów głównego zbioru. Kryterium zakwalifikowania do konkretnego zbioru jest osiągnięcie jak największego podobieństwa elementów wewnątrz zbiorów i jednocześnie maksymalizacja różnic pomiędzy tymi zbiorami. Zdarza się, że elementy zbioru znajdują się tak blisko siebie, że jednoznaczne ich zakwalifikowanie do jednej z grup mogłoby fałszować wyniki. Wtedy sprawdzają się algorytmy grupowania oparte na zbiorach rozmytych. Jednak, choć zwiększają one dokładność grupowania, to powodują także pewne problemy w interpretacji wyników. Dzieje się tak zwłaszcza wtedy, gdy pojedynczy przykład należy do wielu podzbiorów w zbliżonym stopniu i trudno jest ustalić właściwą funkcję przynależności, a niekiedy jest to wręcz niemożliwe.
4. Algorytmy ewolucyjne, zwane też genetycznymi, czyli naśladowanie procesu doboru naturalnego w przyrodzie. Algorytmy te zakładają, że zamiast od jednego, zaczynamy od wielu propozycji rozwiązań, czyli populacji. Każda z jednostek w populacji jest mniej lub bardziej satysfakcjonującym rozwiązaniem problemu. Stopień satysfakcji mierzymy, podstawiając osobnika z populacji do funkcji celu. Następnym krokiem jest poddanie populacji presji zewnętrznej. W rzeczywistym świecie jest to np. ograniczona ilość dostępnego pożywienia, w przypadku algorytmów natomiast moduł, który eliminuje jednostki mające najniższy rezultat w funkcji celu i jednocześnie promuje te ze stopniem najwyższym. Elementy, które „przeżyją” (osiągają odpowiednią wartość funkcji celu), poddawane są krzyżowaniu i mutacji. Krzyżowanie polega na stworzeniu nowego egzemplarza poprzez losowe przemieszanie

² Dużo uwagi statystycy poświęcają obecnie dziedzinie, która stoi na pograniczu rachunku prawdopodobieństwa i analizy statystycznej – sieciom i kwantyfikacjom Bayesa – dostrzegając w nich znaczny potencjał rozwojowy.

części z dwóch innych, mutacja natomiast jest losową zmianą części z parametrów opisujących egzemplarz. Najłatwiej przedstawić to na rysunku (rysunek 2).



Rys. 2. Schemat operacyjny funkcjonowania algorytmów genetycznych
Źródło: opracowanie własne.

Schemat z rysunku 2 wyraźnie ukazuje procesy kluczowe dla efektywnego funkcjonowania algorytmu genetycznego. Jednak nawet najlepsze rozwiązania krzyżowania, wzmacniania czy najbardziej trafna funkcja celu nie gwarantują sukcesu przetwarzania ewolucyjnego. Niezwykle istotne jest przygotowanie populacji początkowej, a także stworzenie ograniczeń, co do liczby jednostek.

- Reguły logiczne są metodą, która wydaje się być najbardziej naturalna dla zapisywania wiedzy dziedzinowej, a ma zostać wykorzystana w algorytmie komputerowym. Reguły praktycznie zawsze zbudowane są z dwóch części: testu logicznego i konkluzji **IF przesłanka THEN konkluzja**. *Przesłanka* jest testem logicznym złożonym z jednego warunku, ale w niektórych przypadkach może to być także kilka, a nawet kilkanaście warunków połączonych operatorami logicznymi. Zmienne, których wartość ustalana jest jako *konkluzja*, mogą w kolejnych regułach stawać się podstawą wnioskowania czy też częścią testu logicznego, dowolnie komplikując model wnioskowania. Zastosowanie podejścia rozmytego pozwala na stosunkowo proste przekształcenie wiedzy dziedzinowej wyrażonej przez ekspertów w języku naturalnym w bazę reguł logicznych przy użyciu pojęć lingwistycznych, opisanych funkcjami przynależności. W realizacji praktycznych systemów często korzysta się z możliwości oznaczenia reguły stopniem jej pewności, najczęściej z przedziału [0,1]. Tak też dzieje się w naszym artykule. Można to zapisać na przykład tak:

IF *wiedza_studenta* = *duża* **THEN** *ocena* = *wysoka*

ŹRÓDŁOWA BAZA DANYCH

Źródłowa baza danych, na podstawie której stworzono hurtownię danych, wykonana została w MS Access 2003 i zawiera klasyczne sprawozdania finansowe firm. Strukturę danych tak zorganizowano, aby umożliwić proste wprowadzanie danych pozyskanych z bilansów firm.

Składa się ona z czterech tabel zatytułowanych: AKTYWA, PASYWA, RACHUNEK i SPOLKI. Każda z tabel zawiera informacje przypisane do odpowiedniego działu w bilansie, a także pola UWAGI, OPIS, ROK i ID_SPOLKI. Znaczenie poszczególnych pól jest następujące:

- UWAGI – informacje dodatkowe, nie występujące w polach bilansowych, często wskazujące na niepewność lub niedokładność danych, np. „Są wątpliwości co do obciążenia wyniku finansowego”.
- OPIS – nazwa firmy wraz z typem spółki i okresem, który obejmuje sprawozdanie, opisana jako tekst, np.: „ABC-POLAND Spółka z o.o. / 71–000 Szczecin/1998”.
- ROK – rok, którego dotyczy raport.
- ID_SPOLKI – identyfikator pozwalający połączyć dany bilans z przechowywaną w tabeli SPOLKI nazwą firmy.

Poza przedstawionymi powyżej wspólnymi polami, w nazwie i funkcji, choć już nie w zawartości, poszczególne tabele zawierają następujące pola³:

Tabela AKTYWA

- Majątek trwały
 - Rzeczowy majątek trwały
 - Finansowy majątek trwały
 - Należności długoterminowe
- Majątek obrotowy
 - Zapasy
 - Należności i roszczenia
 - Papiery wartościowe
 - Środki pieniężne
- Rozliczenie międzyokresowe kosztów
- Suma aktywów

Tabela PASYWA

- Kapitał własny
 - Kapitał podstawowy
 - Należne wkłady na kapitał podstawowy
 - Kapitał zapasowy
 - Niepodzielony wynik finansowy
- Rezerwy
 - Rezerwy na podatek dochodowy
 - Pozostałe rezerwy
- Zobowiązania długoterminowe
 - Długoterminowe pożyczki, papiery wartościowe

³ W tym miejscu podano jedynie podstawowe pola. W rzeczywistości liczba pól jest większa, bowiem definiują one wszystkie pozycje odpowiednich sprawozdań finansowych, które wymagane są prawem. Nazwy pól podawane są małymi literami, ponieważ nie są to identyfikatory, jakie występują w bazie danych, a tylko ich znaczeniowa interpretacja. Prawdziwe, techniczne nazwy skonstruowane są przy użyciu liter, które opisują dział, i liczb do oznaczenia kolejnej pozycji w dziale.

- Długoterminowe kredyty bankowe
- Pozostałe zobowiązania długoterminowe
- Zobowiązania krótkoterminowe i fundusze specjalne
 - Zobowiązania krótkoterminowe
 - Fundusze specjalne
- Rozliczenia międzyokresowe kosztów i pozostałe przychody operacyjne
 - Bierne rozliczenia międzyokresowe kosztów
 - Pozostałe przychody operacyjne
- Suma pasywów

Tabela RACHUNEK

- Przychody ze sprzedaży
- Koszty działalności operacyjnej
- Zysk ze sprzedaży
- Pozostałe przychody operacyjne
 - Ze sprzedaży majątku trwałego
 - Dotacje
 - Pozostałe przychody operacyjne
- Pozostałe koszty operacyjne
 - Wartość sprzedaży składników majątku trwałego
 - Pozostałe koszty operacyjne
- Zysk na działalności operacyjnej
- Przychody finansowe
 - Dywidendy z tytułu udziałów
 - Odsetki uzyskane
 - Pozostałe
- Koszty finansowe
 - Odpisy aktualizujące majątek finansowy
 - Odsetki do zapłacenia
 - Pozostałe
- Zysk z działalności gospodarczej
- Zyski nadzwyczajne
- Straty nadzwyczajne
- Zysk brutto
- Obciążenia wyniku finansowego
 - Podatek dochodowy
 - Pozostałe obciążenia
- Zysk netto

Tabela SPOLKI

- Nazwa spółki
- Miejscowość

– Stan (**F** – funkcjonuje, **U** – w upadłości).

Zastosowanie formatu liczbowego dla opisu pozycji bilansu oznacz jednak brak informacji o ewentualnej niekompletności jakiegoś sprawozdania. „0” w pozycji „Dotacje” może świadczyć zarówno o braku dotacji, jak też o braku informacji o wielkości dotacji. Tabele AKTYWA, PASYWA i RACHUNEK zawierają po 110 rekordów, tabela SPOLKI składa się z 55 rekordów⁴.

Istnieje wiele sposobów tworzenia modeli, które można tu wykorzystać. Jedną z możliwości jest wykorzystanie wiedzy ekspertów, którzy na podstawie swojego doświadczenia zbudują funkcje przynależności dla poszczególnych wymiarów. Inną metodą jest tworzenie funkcji przynależności, opierając się na statystykach (histogramy) etykietowanych przypadków odpowiadających modelowi.

Ta druga metoda wydaje się być celowa do tworzenia modeli przedsiębiorstw „bankrutujących” i „funkcjonujących” ze względu na występowanie w bazie danych etykietowanych przypadków. Pierwsza metoda może być wykorzystana do tworzenia modelu np. idealnego przedsiębiorstwa.

OCENA JAKOŚCI TWORZONEGO MODELU

Jeśli potrafimy zbudować model, to należy zastanowić się nad oceną jego jakości. Ocena jakości tworzonego modelu może odbywać się co najmniej na dwóch płaszczyznach. Pierwsza to dokładność klasyfikacji obiektów, tzn. ile obiektów mających taką samą etykietę jak konkretny model zostaje sklasyfikowanych jako bardzo do niego podobne. Istnieje tu niebezpieczeństwo popełnienia błędu statystycznego drugiego rodzaju, który polega na zakwalifikowaniu obiektu o odmiennej etykiecie, jako podobnego do modelu wzorca. Dążenie do zwiększenia praktycznej użyteczności i dokładności modelu powinno być oceniane zawsze jako minimalizacja błędów obu rodzajów (zarówno pierwszego, jak i drugiego).

Inny sposób pomiaru jakości modelu wynika wprost z właściwości zbiorów rozmytych użytych do konstrukcji modelu. Wśród ogólnych cech właściwego modelu bierzemy pod uwagę przede wszystkim:

– Możliwie niewielką szerokość zbioru rozmytego w stosunku do przyjmowanych wartości danej cechy (selektywność modelu):

$$\alpha = \frac{|\text{Sup } p\Theta|}{\lambda} \quad (1)$$

gdzie:

λ – szerokość przedziału przyjmowanych wartości,

$\text{Sup } p()$ – nośnik zbioru rozmytego; zbiór elementów, dla których wartość funkcji przynależności jest większa od zera.

– Stopień rozmycia danego zbioru; rozmycie modelu jest sumą rozmyć jego poszczególnych wymiarów:

⁴ Wynika to z prostego faktu, że bazę zbudowano z 55 sprawozdań, ale każde z tych sprawozdań zawierało dane z dwóch kolejnych lat, co dało 110 rekordów w pozostałych tablicach.

$$\beta = \frac{|\text{Core}(\Theta)|}{|\text{Sup } p\Theta|} \quad (2)$$

gdzie:

$\text{Core}()$ – rdzeń zbioru rozmytego, czyli zbiór takich elementów, dla których wartość funkcji przynależności jest równa 1.

– Istnienie co najmniej jednego takiego punktu w przestrzeni cech, który całkowicie spełnia model (realność modelu):

$$\forall x \in [1, 2, \dots, n], \text{Height}(M_{A^x}) = 1 \quad (3)$$

gdzie:

n – liczba wymiarów modelu,

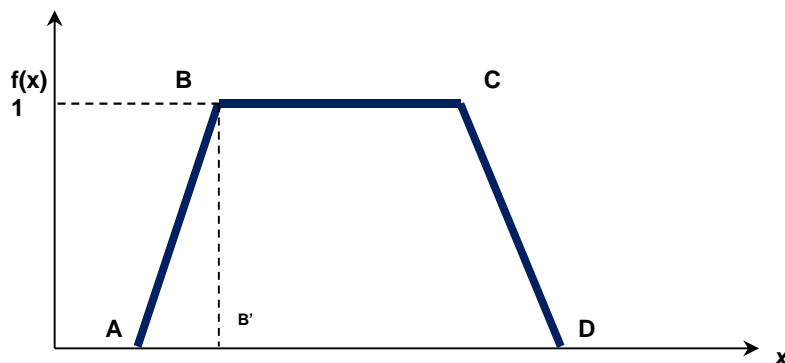
M_A – konkretny model,

$\text{Height}()$ – wysokość zbioru rozmytego; największa wartość funkcji przynależności; zbiór rozmyty jest znormalizowany wtedy i tylko wtedy, gdy $h = 1$.

Do opisu przykładowych modeli, pokazanych poniżej jako znormalizowane funkcje przynależności, zastosowano funkcje trapezowe. Postać trapezowa funkcji przynależności wynika z faktu, że rozkład poszczególnych wskaźników jest jednomodalny, a także ze względu na bardzo praktyczne właściwości funkcji trapezowych.

Przez zmniejszenie $\text{Core}()$ do punktu przekształcić można funkcję trapezową w trójkątną funkcję przynależności, a tę z kolei bez większego problemu przeliczyć z dowolną dokładnością na funkcję sigmoidalną. Wydaje się więc, że taki wybór daje największą elastyczność przy jednoczesnym uproszczeniu praktycznej realizacji.

Wymagania dane wzorami (1), (2) i (3) mają charakter rozbieżny, ale konstrukcję zbioru rozmytego rozwiązać można jako zadanie optymalizacji. Dodatkowo funkcja trapezowa opisana jest przez kilka parametrów, a ich konstrukcja jest dość łatwa i w pewnym stopniu intuicyjna. Na rysunku 2 przedstawiono trapezową funkcję przynależności wraz z opisującymi ją parametrami. Przyjęta przy opisie modeli funkcji przynależności konwencja zapisu to $\{A; B; C; D\}$.



Rys. 2. Sposób opisu znormalizowanej trapezowej funkcji przynależności

Budowa części wzrastającej i opadającej funkcji przynależności prowadzona może być niezależnie, ale algorytm rozwiązania jest jednakowy. Przyjmując, że jakość reprezentacji danych numerycznych przez zbiór rozmyty Θ może być oceniana na podstawie sumy stopni przynależności, to jakość taką opisać można jako żądanie maksymalizacji zależności:

$$\max \sum_k \theta[x(k)] \quad (4)$$

W przypadku wznoszącej się części funkcji przynależności poszukujemy takiej wartości ograniczenia dolnego (punkt A na rysunku, dla którego wartość funkcji przynależności osiąga wartość równą zero), dla której suma ze wzoru (4) osiąga wartość maksymalną. Tak rozumiany wskaźnik jakości osiąga coraz wyższe wartości dla coraz niższych wartości odciętej punktu A.

Drugie wymaganie zwartości zbioru rozmytego sformułujemy w postaci minimalizacji jego nośnika, czyli według poniższej zależności (przy czym B' jest rzutem punktu B na oś x):

$$\min |AB'| \quad (5)$$

Zadanie maksymalizacji kryterium (4) i minimalizacji kryterium (5) można przedstawić wspólnie jako zadanie maksymalizacji wyrażenia (6):

$$\max_{AB' \neq 0} \frac{\sum_k \theta[x(k)]}{|AB'|} \quad (6)$$

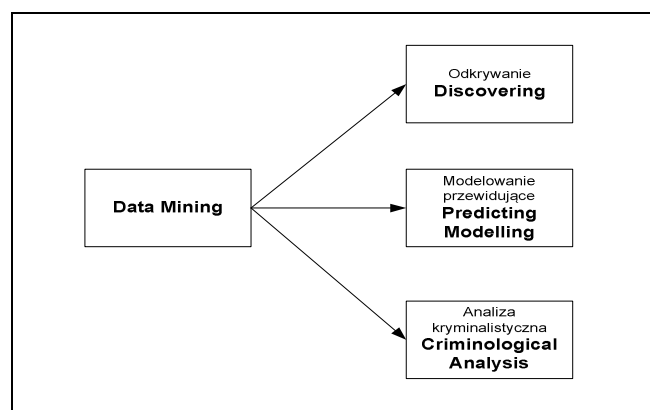
Konstrukcja odpowiedniej postaci opadającej części funkcji przynależności jest realizowana w sposób analogiczny. Konstrukcja rdzenia zbioru rozmytego oparta jest na medianie, co wpływa korzystnie na poprawę odporności procesu na zakłócenia i błędy. Rdzeń zbioru rozmytego może być albo jedną wartością (dla nieparzystej liczby próbek w oknie granulacji), albo przedziałem numerycznym (w przypadku parzystej liczby próbek).

W praktyce niewdzięcznym zadaniem jest opracowanie algorytmu identyfikacji funkcji przynależności. Zaproponowana tu metoda identyfikacji jest o tyle interesująca, że nie wymaga badania właściwości rozkładu prawdopodobieństwa w zbiorze i jest możliwa do stosowania bez żadnych ograniczeń. W pracy obliczenia wykonano metodą enumeratywnego poszukiwania optymalnego rozwiązania, bowiem badany zbiór nie był na tyle liczny, aby konieczne było stosowanie zorganizowanych metod poszukiwania ekstremum.

BADANIA PRZYDATNOŚCI PROPONOWANEJ METODY DO OPISU BANKRUTUJĄCYCH I FUNKCJONUJĄCYCH PRZEDSIĘBIORSTW

Powyżej pokazano założenia metody służącej ekstrakcji wiedzy z hurtowni danych, jako naturalnego źródła informacji dla systemów wspomaganie decyzji w przedsiębiorstwie.

W praktycznych zastosowaniach Data Mining'u zaobserwować można trzy główne nurty różniące się podejściem i celami. Strukturę podziału przedstawiono na rysunku 3.

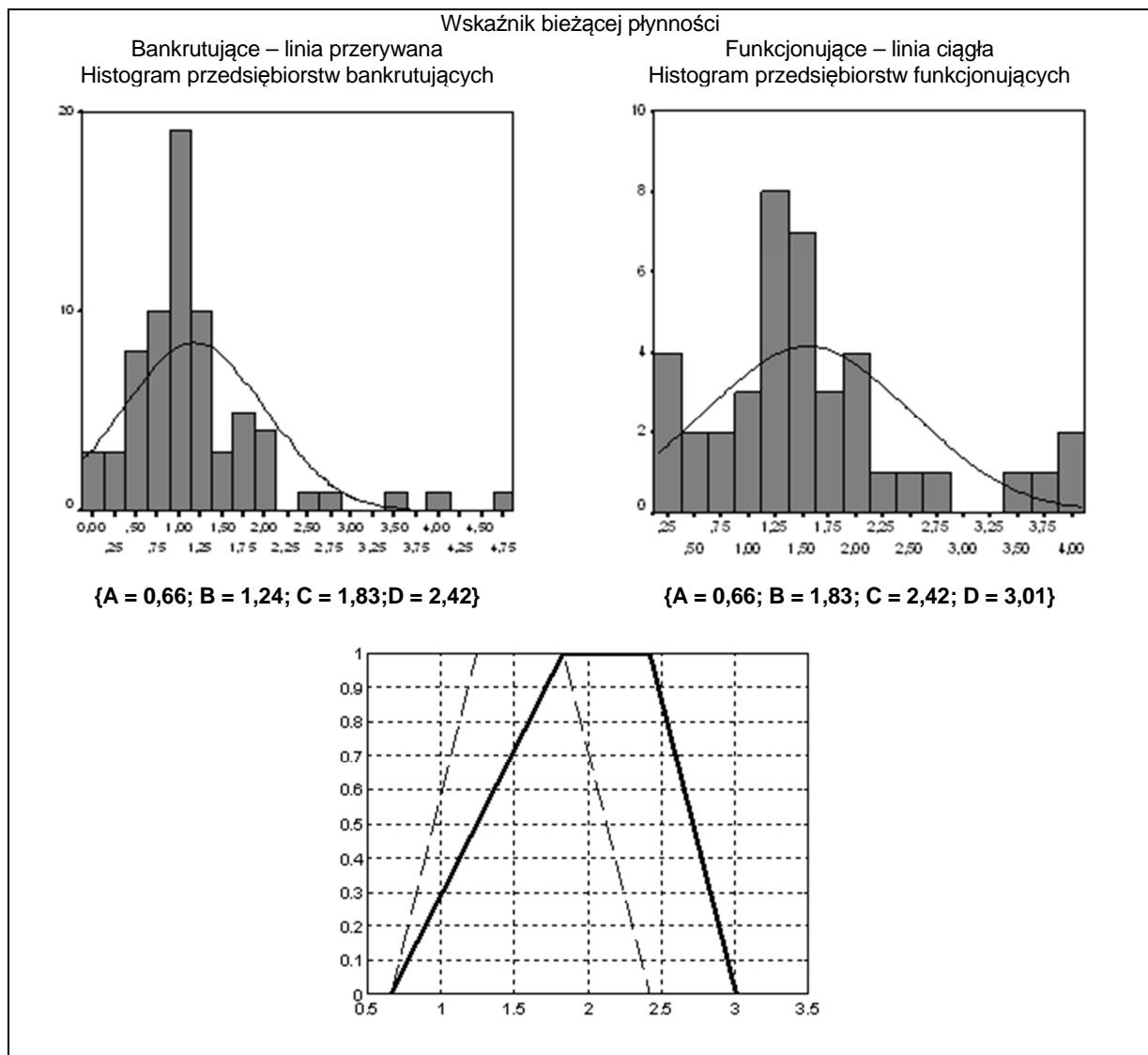


Rys. 3. Główne obszary zastosowań Data Mining'u
Źródło: opracowanie własne na podstawie Ramachandran (2001).

Górna gałąź, **odkrywanie**, odpowiada za znajdowanie w dostępnych danych wartościowych powiązań pozwalających objaśniać przebieg i korelację zjawisk, które miały już miejsce. **Modelowanie przewidujące** skupia się na próbach określenia przyszłych stanów systemu (przedsiębiorstwa, osoby chorej, urzędnika) na podstawie danych historycznych. Ostatnia gałąź podziału, czyli **analiza kryminalistyczna**, jest połączeniem matematycznych metod ekstrakcji danych, z rezultatami badań socjologicznych i psychologicznych dotyczących zachowań ludzi (grup społecznych).

Program eksperymentów zainspirowany został pracą (Heine 2000), w której odnajdujemy analizę zdolności dyskryminacyjnej 28 wskaźników finansowych, bazującej na typowo statystycznym podejściu (*F*-test).

Eksperymenty wykonano dla wielu wskaźników, a na rysunku 4 przedstawiono wyniki dla trzech wybranych wskaźników.

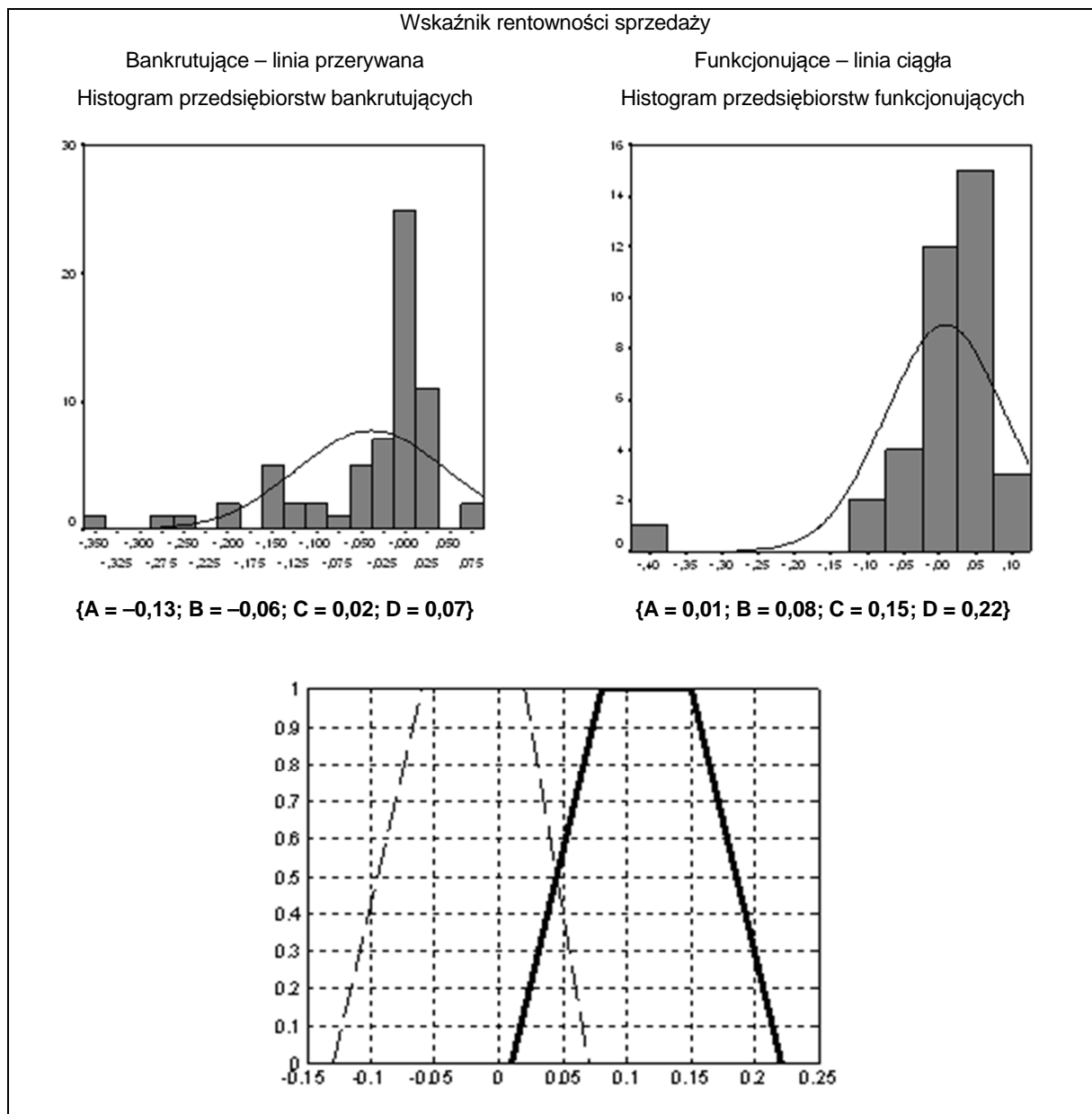


Rys. 4. Analiza statystyczna (histogram) wskaźnika bieżącej płynności dla przedsiębiorstw bankrutujących i funkcjonujących oraz funkcje przynależności dla obu zbiorów przedsiębiorstw
Źródło: opracowanie własne na podstawie Nielek (2004).

Uwaga: oznaczenie wierzchołków trapezu zgodne z rysunkiem 2.

Z rysunku 4 wynika, że wartość dyskryminacyjna bardzo popularnego wskaźnika bieżącej płynności jest dla badanej próby przedsiębiorstw dość ograniczona, co widać wyraźnie, bowiem część wspólna funkcji przynależności dla obu grup przedsiębiorstw jest duża, dlatego etykietowanie prowadzone na podstawie tego wskaźnika może być obarczone sporym błędem.

Na rysunku 5 pokazano analogiczną analizę przeprowadzoną dla wskaźnika rentowności sprzedaży.



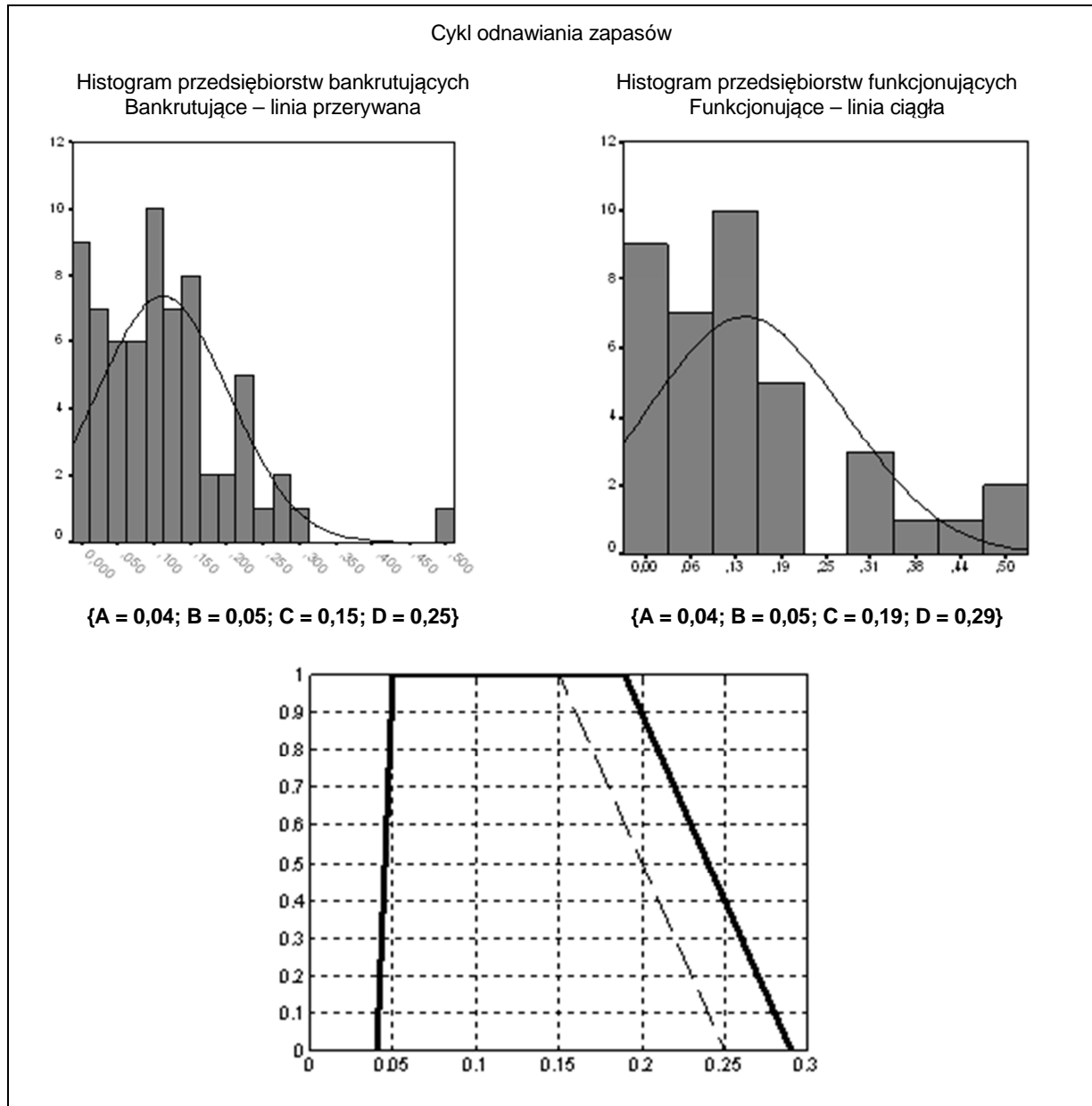
Rys. 5. Analiza statystyczna (histogram) wskaźnika rentowności sprzedaży dla przedsiębiorstw bankrutujących i funkcjonujących oraz funkcje przynależności dla obu zbiorów przedsiębiorstw
Źródło: opracowanie własne na podstawie Nielek (2004).

Uwaga: oznaczenie wierzchołków trapezu zgodne z rysunkiem 2.

Z rysunku 5 wypływa wniosek, że wartość dyskryminacyjna wskaźnika rentowności sprzedaży jest dla badanej próby przedsiębiorstw bardzo wysoka, co wynika z faktu, że część

wspólna funkcji przynależności dla obu grup przedsiębiorstw jest bardzo mała, dlatego etykietowanie prowadzone na podstawie tego wskaźnika powinno być bardzo skuteczne.

Na kolejnym rysunku 6 pokazano następnny wskaźnik, jakim jest zależność opisująca cykl odnawiania zapasów w firmach funkcjonujących i bankrutujących.



Rys. 6. Analiza statystyczna (histogram) wskaźnika cyklu odnawiania zapasów dla przedsiębiorstw bankrutujących i funkcjonujących oraz funkcje przynależności dla obu zbiorów przedsiębiorstw
Źródło: opracowanie własne na podstawie Nielek (2004).

Uwaga: oznaczenie wierzchołków trapezu zgodne z rysunkiem 2.

PODSUMOWANIE

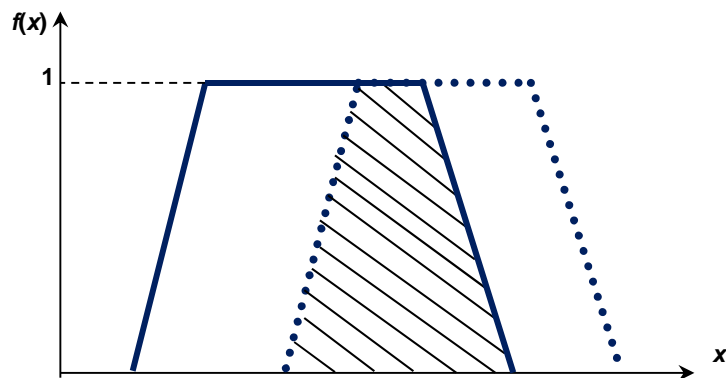
W ramach prowadzonych od lat prac stworzona została rozwijana do dzisiaj aplikacja i interpreter, umożliwiające eksploatację i ocenę zaproponowanego rozwiązania w rzeczywistej hurtowni danych finansowych firm, co nie jest jednak tematem tego artykułu (Nielek 2004).

W prezentowanej metodzie zastosowano podejście rozmyte, czyli odmienne od obecnie powszechnie stosowanego. Przez to możliwe było znaczne uproszczenie rozważań dotyczących oceny zdolności dyskryminacyjnej wskaźników, sprowadzając czynność tę do czysto geometrycznej analizy iloczynu dwóch zbiorów w kształcie trapezu.

Z przeprowadzonych badań wynika, że cel niniejszego artykułu, jakim było zaprezentowanie nietradycyjnego podejścia do oceny zdolności dyskryminacyjnej wskaźników analizy finansowej, czyli rozważenie ich potencjału, jako atrybutów zadania dyskryminacji, został osiągnięty. Zaprezentowana metoda (szczegółowe wyjaśnienie na rysunku 7) umożliwia prostą analizę przydatności do zadań wskaźników finansowych przedsiębiorstwa.

Kontekst problemowy prezentowanej metody polega na tym, że skomplikowaną analizę statystyczną, prowadzącą do określenia zdolności dyskryminacyjnej, zastępujemy znacznie prostszą i równie skuteczną metodą, której zasada sprowadza się do wyboru takich wskaźników, dla których pole wspólne funkcji przynależności (część zakreskowana na rysunku 7) jest najmniejsze. Oznacza to, że wielkość pola wspólnego funkcji przynależności dla przedsiębiorstw funkcjonujących i bankrutujących, charakterystycznego dla określonego wskaźnika, jednoznacznie definiuje przydatność tego wskaźnika do zastosowań w systemach wspomaganie decyzji, których konstrukcja jest bardzo prosta.

Przy okazji, z powyższej dyskusji wynika dość oczywisty związek pomiędzy rozmyciem zbiorów a niepewnością klasyfikacji. Możemy związek ten przedstawić następująco: *im większy stopień rozmycia zbioru, tym więcej informacji niesie, a więc mniejsza jest niepewność*. Jako miarę niepewności zbioru rozmytego możemy przyjąć miarę jego ostrości.



Rys. 7. Prosty sposób wyznaczenia zdolności dyskryminacyjnej z wykorzystaniem funkcji przynależności
Źródło: opracowanie własne.

Największa niepewność występuje w przypadku, kiedy porównywane przedziały rozmyte są dość podobne, choć niekoniecznie identyczne (rysunek 6). W przypadku, kiedy porównujemy identyczne przedziały lub jeden z nich jest większy od drugiego (nie mają części wspólnej), niepewność jest zerowa. W przypadku, kiedy jeden przedział jest zdecydowanie większy od drugiego, ale mają pewną niewielką część wspólną, niepewność wyniku również nie jest duża (rysunek 5).

Postać uzyskanego systemu wspomaganie decyzji zależy oczywiście nie tylko od przyjętej techniki ekstrakcji wiedzy, ale również od zastosowanej procedury przetwarzania danych, zadań stawianych systemowi wspomaganie decyzji i od merytorycznych aspektów wiedzy dziedzinowej.

Prezentowany materiał nie rozstrzyga o przydatności konkretnych wskaźników, prezentując jedynie metodę badawczą, która prosto i jednoznacznie wskazać może właściwy podzbiór wskaźników, które zastosowane być mogą do systemu informacyjnego przedsiębiorstwa.

PIŚMIENNICTWO

- Heine M.L.** Predicting Financial Distress of Companies: Revisiting the Z-Score and ZETA Models; This paper is adapted and updated from E. Altman, „Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy”, J. Finance, September 1968; and E. Altman, R. Haldeman and P. Narayanan, „Zeta Analysis: A New Model to Identify Bankruptcy Risk of Corporations”, J. Banking & Finance, 1, 1977; Stern School of Business, New York University, July 2000 (dostępne na prawach rękopisu).
- Nielek R.** 2004. Możliwości definicji składni języka zapytań bazy danych w przedsiębiorstwie; Praca dyplomowa przygotowana pod kierunkiem L. Rozenberga. Szczecin.
- Pushpa Ramachandran M.** 2001. Mining for Gold, Wipro Technologies, www.wipro.co
- Rozenberg L.** 2000. Analiza wskaźnikowa firm o różnych strukturach bilansu, Firma i Rynek, nr 01, 64–67; Szczecin.
- Śmiałkowska B.** 2009. Metoda dopasowania hurtowni danych do zmiennych potrzeb informacyjnych przedsiębiorstwa, rozprawa habilitacyjna; Wydaw. Uczelniane Zachodniopolskiego Uniwersytetu Technologicznego w Szczecinie, ISBN 978-8-376-63006-9.