

KLASYFIKACJA METOD PROGNOZOWANIA STANU ZAGROŻENIA UPADŁOŚCIĄ PRZEDSIĘBIORSTW

Jan Andreasik

Wyższa Szkoła Zarządzania i Administracji w Zamościu

Celem artykułu jest przegląd i klasyfikacja metod prognozowania zagrożenia upadłością przedsiębiorstw, które mogą posłużyć do budowy systemu wczesnego ostrzegania przed zagrożeniem ciągłości działalności małych i średnich przedsiębiorstw województw lubelskiego i podkarpackiego.

Próbę klasyfikacji metod prognozowania stanu zagrożenia upadłością przedsiębiorstw podjęto na wydziałach ekonomicznych uniwersytetu w Gencie (Belgia) oraz uniwersytetu Loughborough (Wielka Brytania). S. Balcaen, H. Ooghe¹ przedstawili klasyfikację metod statystycznych w oparciu o przegląd literatury z okresu od 1967 roku (pierwsza praca Beavera) do 2004 roku. Ci sami autorzy w innym raporcie² przedstawili klasyfikację metod alternatywnych z zakresu sztucznej inteligencji oraz metodologii wielokryterialnego podejmowania decyzji. M. Aziz i H. A. Dar³ przedstawili stan badań dotyczących prognozowania stanu zagrożenia upadłością dzieląc modele na trzy grupy: statystyczne, modele sztucznej inteligencji, modele oparte o inne teorie.

Autor niniejszego opracowania przeprowadził obszerną analizę literatury, w wyniku której dokonał następującej klasyfikacji metod prognozowania stanu zagrożenia upadłością przedsiębiorstw:

A. Metody statystyczne:

1. Modele regresji wielorakiej:
 - modele logitowe,
 - modele probitowe
2. Modele analizy wielowymiarowej:
 - analiza dyskryminacyjna (discriminant analysis)
 - analiza skupisk (cluster analysis),
 - skalowanie wielowymiarowe (multidimensional scaling).

¹ Balcaen S, Ooghe H. (2004) 35 years of studies on business failure: an overview of the classical statistical methodologies and their related problems, Universiteit Gent, Faculteit Economie, June 2004, 2004/248, www.vlerick.be/research/workingpapers/vlgms-wp-2040-15.pdf

² Balcaen S., Ooghe H. (2004) Alternative methodologies in studies on business failure: do they produce Belter results than the classic statistical methods?, Vleric Leuven Gent Working Paper Series 2004/16, www.vlerick.be/research/workingpapers/vlgms-wp-2040-16.pdf

³ Aziz M.A., Dar H.A. (2004) Predicting Corporate Bankruptcy: Whether do we stand? Department of Economics, Loughborough University, UK, <http://gnu.univ.gda.pl/~eefs/pap/aziz.doc>

B. Metody sztucznej inteligencji:

1. Metody maszynowego uczenia (machine learning)
 - metody klasyfikacji oparte na algorytmach generowania drzew decyzyjnych (CHAID Chi-square automatic interaction detection), CART Classification and Regression Trees, C4.5 Information Theory),
 - sieci przekonań Bayesa,
 - metody rozumowania na podstawie analizy serii przypadków (Case-Based Reasoning),
 - metody generowania reguł wnioskowania.
2. Techniki wykorzystywane w metodach klasyfikacji:
 - sieci neuronowe,
 - algorytmy bazujące na teorii zbiorów rozmytych,
 - algorytmy genetyczne,
 - algorytmy analizy syntaktycznej,
 - algorytmy bazujące na teorii zbiorów przybliżonych.

C. Metody wielokryterialnego podejmowania decyzji:

1. Sortowanie metodą ELECTRE TRI,
2. Sortowanie z zastosowaniem funkcji użyteczności UTADIS (UTilites Additives DIScriminantes).

D. Modele symulacyjne (symulatory tworzone są w oparciu o teorię entropii, teorię gier, teorię katastrof, teorię chaosu).

E. Standardowe wskaźniki agencji ratingowych (np. wskaźnik KMV agencji Moody's).

Podstawą powyższej klasyfikacji są teorie wykorzystywane do budowy modeli. Można tu wymienić teorię automatycznej klasyfikacji obiektów, teorię wielokryterialnego podejmowania decyzji⁴, teorię rozumowania opartą na analizie przypadków (Case-Based Reasoning)⁵, teorię maszynowego uczenia.

W rozpoznawaniu obiektów wyróżnia się dwa podejścia różniące się sposobem traktowania rozpoznawanego obiektu.

1. Podejście decyzyjno-teoretyczne, w którym każdy obiekt reprezentowany jest przez wektor cech widziany jako punkt w przestrzeni cech.
2. Podejście strukturalne, w którym każdy obiekt jest reprezentowany jako złożenie składowych pierwotnych za pomocą różnych relacji, jakie mogą zachodzić między składowymi. Dodatkowo, występuje pojęcie tzw. modelu (prototypu) klasy, który zawiera charakterystykę wszystkich obiektów tworzących daną klasę.

Obszerną analizę algorytmów automatycznej klasyfikacji obiektów zawiera praca K. Stapor⁶.

⁴ Zopounidis C. Dimitras A.I. (1998) Multicriteria Decision Aid Methods for the Prediction of Business Failure. Kluwer Academic Publishers.

⁵ Pal S.K., Shiu S.C.K. (2004) Foundations of soft case-based reasoning. John Wiley & Sons. Inc.

⁶ Stapor K., (2005) Automatyczna klasyfikacja obiektów. Akademia Oficyna Wydawnicza EXIT, Warszawa.

Inny układ klasyfikacyjny można zbudować na podstawie zasadniczych cech modelu. Autor niniejszej pracy proponuje układ pięciu cech charakteryzujących model prognozowania:

Warunek prognozowania	Charakterystyka obiektu	Cechy zbioru uczącego	Sposób wyjaśniania	Charakterystyka klas
-----------------------	-------------------------	-----------------------	--------------------	----------------------

Na podstawie dotychczasowych modeli prognozowania można określić następujące warunki prognozowania (zaliczania obiektu do danej klasy):

- scoring, (wyznaczenie punktu granicznego cut-off point)
- prawdopodobieństwo zaliczenia do danej klasy,
- czas do upadłości,
- indeks zgodności ze wzorcem (prototypem),
- przynależność do języka.

Model charakteryzowany jest przez układ zmiennych objaśniających:

- uznane za istotne zmienne objaśniające,
- wszystkie podane zmienne objaśniające,
- kryteria decyzyjne,
- struktury symboliczne (ciąg, drzewo, graf, struktura relacyjna),
- gramatykę.

Zbiór uczący może spełniać następujące warunki:

- warunki statystyczne,
- niepełne dane
- niejednoznaczne przyporządkowanie.

W dotychczasowych modelach spotyka się następujące sposoby wyjaśniania:

- drzewa klasyfikacyjne,
- reguły warunkowe,
- gramatyki,
- wzorce,
- prawdopodobieństwa warunkowe.

Zwyczaj w modelach stosowanych do prognozy stanu zagrożenia stosuje się jedną zmienną objaśnianą. Jest nią zmienna wskazująca przynależność danego przedsiębiorstwa określonego wektorem wskaźników lub cech do dwóch klas: do grupy przedsiębiorstw zagrożonych upadłością oraz do grupy przedsiębiorstw nie zagrożonych. Z uwagi na trudności z jednoznacznym zdefiniowaniem stanu upadłości przedsiębiorstwa wyróżniane są inne zmienne np. populacja grup produkcyjnych, w których odnotowano zysk netto oraz, w których odnotowano stratę netto. W pracy J. Khalil, J. Martel, P. Jurtas⁷ wyróżniono pięć klas ryzyka C1 - ryzyko, C2 - wysokie ryzyko, C3 - ryzyko akceptowane, C4 - normalne ryzyko, C4 - małe ryzyko.

Pierwszą kategorią modeli są modele opracowane na gruncie statystyki matematycznej.

Najbardziej znaną i jednocześnie najstarszą metodą prognozowania stanu zagrożenia upadłością przedsiębiorstw jest analiza dyskryminacyjna zastosowana przez Edwarda Altmana profesora New York University⁸. Metoda ta jest wykorzystywana do tworzenia modeli prognozowania w wielu krajach. C.Y. Shirata przeprowadził badania w Japonii⁹ na grupie 685 firm poddanych likwidacji i 300 firm o dobrej kondycji. O. Hajdu, M. Virag przeprowadzili podobne badania na Węgrzech¹⁰. W Polsce badania przy użyciu tej metody przeprowadziła Elżbieta Maczyńska¹¹ w INE PAN analizując 40 firm zagrożonych i 40 firm nie zagrożonych upadłością notowanych na giełdzie warszawskiej. Podobne badania przeprowadziły Dorota Appenzeller i Katarzyna Szarzec¹² również dokonując analizy spółek giełdowych. W badaniach Doroty Kwiatkowskiej-Ciotucha i Urszuli Załuskiej¹³ z Akademii Ekonomicznej we Wrocławiu przyjęto klasyfikację na dwie grupy przedsiębiorstw: pierwszą, w której odnotowano zysk netto, i drugą, w której odnotowano stratę netto wg danych GUS dla grup przedsiębiorstw produkcyjnych. Do oszacowania parametrów funkcji dyskryminacyjnej wykorzystano pakiet STATISTICA.

Istotą analizy dyskryminacyjnej jest podział (klasyfikacja) zbioru na grupy. W analizie dyskryminacyjnej dąży się do utworzenia kombinacji liniowej zmiennych niezależnych, która najlepiej dyskryminuje (rozdziela) dwie (lub więcej) grupy określone „a priori”¹⁴.

Głównym zarzutem, który jest formułowany odnośnie stosowania tej metody jest to, że wskaźniki finansowe zastosowane jako zmienne objaśniające użyte są celem maksymalizacji funkcji dyskryminacyjnej. Powstaje model klasyfikacyjny a nie prognostyczny. Zatem nie jest celowe rozważanie relacji przyczynowo-skutkowej pomiędzy uwzględnionymi w modelu wskaźnikami a stanem zagrożenia upadłością. Utrudnia to właściwe

⁸ Altman E. I. (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *The Journal of Finance*, vol. XXIII, no.4, pp.589-609.

⁹ Shirata C.Y. (1998) Financial ratios as predictors of bankruptcy in Japan: An empirical research. www.bke.hu/vpenzogy/download/japan.pdf

¹⁰ Hajdu O., Virag M. (1991) A Hungarian model for predicting financial bankruptcy, www.lib.uni-covinus.hu/gt/2001-1-2/hajdu-virag.pdf

¹¹ Maczyńska E. (2004) Systemy wczesnego ostrzegania. *Nowe Życie Gospodarcze*, nr 12, 4.

¹² Appenzeller D., Szarzec K., (2004) Prognozowanie zagrożenia upadłością polskich spółek publicznych. *Inżynieria Finansowa*, nr 1, 120-128.

¹³ Kwiatkowska-Ciotucha D., Załuska U., (2003) Możliwości zastosowania analizy dyskryminacyjnej do oceny sytuacji branż produkcyjnych w Polsce. *Prace Naukowe Akademii Ekonomicznej we Wrocławiu, Ekonometria* 12, nr 1002, 137-143.

¹⁴ Aczel Amir D. (2000) *Statystyka w zarządzaniu*. Wydawnictwa Naukowe PWN, Warszawa, 882-900.

⁷ Khalil J., Martel J., Jurtas P. (1999) A multicriterion system for credit risk rating. *Faculte des sciences de l'administration Université Laval Quebec Canada Document de travail 1999-014*, <http://ideas.repec.org/p/fth/lavadm/99-014.html>

postawienie diagnozy w aspekcie analizy przyczyn stanu zagrożenia upadłością danego przedsiębiorstwa. W analizie dyskryminacyjnej wprowadza się na początku listę zmiennych objaśniających, z której do modelu wybranych zostaje kilka najlepiej prognozujących podział na dwie grupy przedsiębiorstw. Musi być zachowany postulat o wzajemnym nie skorelowaniu zmiennych. Istotnym założeniem w tej metodzie jest przyjęcie, że analizowane populacje mają wielowymiarowe rozkłady normalne o równych macierzach wariancji/kowariancji i niekoniecznie równych średnich. Po wyznaczeniu funkcji dyskryminacyjnej, kolejny etap stanowi wyznaczenie wartości granicznej (tzw. cut off point) pozwalającej na zakwalifikowanie określonej jednostki do danej grupy. W tym celu należy wyznaczyć wartości średnie funkcji dyskryminacyjnej dla poszczególnych grup. Przy wyznaczeniu wartości granicznej najczęściej zakłada się, że znajduje się ona w połowie między wartościami średnimi funkcji z poszczególnych grup.

W pierwszym modelu Altmana wartość graniczna wskaźnika cut off point wynosi 2,675. Strefa pośrednia rozciąga się między 1,81 a 2,99. W przypadku Z-score <1,80 zagrożenie upadłością wzrasta. W modelu Kwiatkowskiej-Ciotuchy i Załuskiej wskaźnik cut off wynosi 0,265. W modelu japońskim Shirata ustalił cut off point równy 0,38.

Inną metodą statystyczną stosowaną do określenia stanu zagrożenia upadłością jest regresja wieloraka. Najbardziej znanym modelem jest model logitowy Ohlsona. W tych modelach określa się prawdopodobieństwo tego, że zmienna zależna Y przyjmie wartość 1 (prawdopodobieństwo sukcesu), tzn. nastąpi zaklasyfikowanie do grupy przedsiębiorstw zagrożonych upadłością. W analizie logitowej korzystnymi cechami są: brak założenia o normalności rozkładu poszczególnych zmiennych oraz brak założenia o równości macierzy kowariancji poszczególnych grup. W analizie regresji często wprowadza się zmienną wskaźnikową (sztuczną). Zmienna ta dotyczy określenia jakościowego poziomu A. Zmienna przyjmuje wartość 1 jeżeli poziom A został osiągnięty, oraz wartość 0 jeżeli poziom A nie został osiągnięty. W modelu Ohlsona są ujęte dwie takie zmienne. X5 jest równe 1, gdy suma zadłużenia przedsiębiorstwa jest większa od sumy aktywów oraz jest równe 0 w przeciwnym przypadku. X8 jest równe 1 jeżeli wynik netto jest mniejszy od zera w ostatnich dwóch latach, oraz 0 w przeciwnym przypadku.

Wraz ze wzrostem wartości zmiennej objaśnianej Z wskaźnik Y, czyli prawdopodobieństwo bankructwa rośnie. Z tego wynika, że wzrost poszczególnych wag działa stymulująco na wzrost zagrożenia bankructwem a spadek destymulująco. Wartość wskaźnika Y zawarta jest w przedziale od 0 do 1. Nie oznacza to jednak, że punkt graniczny (cut-off point) dla modelu logitowego wynosi 0,5. Punkt graniczny może być nawet mniejszy od 0,1. Np. w modelu Ohlsona wynosi 0,038.

W modelu probitowym zamiast funkcji logitowej występuje skumulowana funkcja rozkładu normalnego. Badania prognozowania stanu upadłości przedsiębiorstw za pomocą tej metody pierwszy przeprowadził Żmijewski¹⁵. Tseng F., Lin L.¹⁶ wprowadzili model z kwadratową funkcją logitową. W tym modelu optymalny punkt graniczny rozdziałający przedsiębiorstwa zagrożone upadłością oraz nie zagrożone wynosi 0,639.

S.Balcaen i H. Ooghe przedstawili w swojej pracy kilka problemów dotyczących wykorzystania metod statystycznych w prognozowaniu stanu zagrożenia upadłością przedsiębiorstw. Pierwszym problemem jest ustalenie precyzyjnej definicji zmiennej zależnej, która powinna być w modelach statystycznych dychotomiczna. W rzeczywistości trudno jednoznacznie zweryfikować upadłość przedsiębiorstwa. Istotnym jest tu jednak arbitralne określenie warunku, który dzieli zbiór przedsiębiorstw na dwa rozłączne podzbiory. Tylko wówczas można przeprowadzić klasyfikację. Drugim problemem jest dobór próby statystycznej przedsiębiorstw. W każdym przypadku wszystkie zmienne objaśniające muszą być kompletne. Również zbiór „uczący” musi zawierać pary przedsiębiorstw odpowiednio zagrożone i nie zagrożone upadłością. Trzecim problemem jest wymaganie aby zmienne modelu były stabilne w czasie. Czwarty problem to interpretacja zmiennych objaśniających, którymi są wskaźniki finansowe obliczane na podstawie rachunku zysków i strat oraz bilansu. Istnieją trudności w jednoznacznej definicji wskaźników szczególnie gdy zmienia się układ poszczególnych sprawozdań finansowych. Piątym problemem jest wybór niezależnych zmiennych. Do tej pory nie ma teorii zagrożenia upadłością przedsiębiorstw. Zatem ustalenie zestawu zmiennych do analizy jest arbitralną decyzją prowadzących badania. Ostatnim problemem jest uwzględnianie w analizie statystycznej stanu zagrożenia a nie procesu jaki w czasie rzeczywistym przebiega w przedsiębiorstwie. Statyczny klasyczny model statystyczny ignoruje zmiany jakie zachodzą w przedsiębiorstwie co do np. rozmiaru działalności w określonym przedziale czasowym (kilku lat objętych badaniami).

Drugą kategorią modeli są modele opracowane na podstawie teorii maszynowego uczenia. W teorii tej procesy uczenia podlegają następującym paradygmatom¹⁷:

- paradygmat statystyczny,
- paradygmat syntaktyczny,
- paradygmat konekcyjny,
- paradygmat ewolucyjny,

¹⁵ Żmijewski M.E. (1984) Methodological issues related to the estimation of financial distress prediction models. Studies on current Econometric Issues in Accounting Research, 59-82.

¹⁶ Steng Fang-Mei, Lin Lin (2005) A quadratic interval logit model for forecasting bankruptcy. OMEGA 33, 85-91.

¹⁷ Glazer A. P. (2003) Automatyzacja procesu pozyskiwania wiedzy dla systemów ekspertowych. Praca doktorska. Politechnika Gdańska.

- paradygmat analityczny,
- paradygmat indukcyjny.

Według paradygmatu statystycznego budowane są modele wykorzystujące sieci Bayesa¹⁸. Sieci Bayesa stanowią probabilistyczne modele graficzne reprezentujące zbiór zmiennych losowych i relacje probabilistyczne między nimi. Strukturę BN (Bayesian Networks) reprezentuje skierowany acykliczny graf (DAG direct acyclic graph), w którym wierzchołki reprezentują zmienne a łuki wyrażają zależności między zmiennymi. Część probabilistyczna BN jest reprezentowana przez zbiór prawdopodobieństw warunkowych. Prawdopodobieństwo $P(A)$ reprezentuje przekonanie (belief) przed obserwacją danego zdarzenia. $P(A)$ może być miarą tego jaki procent przedsiębiorstw jest zagrożonych upadłością w całej populacji danego regionu. $P(B)$ reprezentuje prawdopodobieństwo wskaźnika upadłości (zmiennej objaśniającej), np. że sprawozdanie finansowe F-02 jest wykonane z dużym opóźnieniem. $P(B|A)$ oznacza prawdopodobieństwo, że zbiór przedsiębiorstw jest zagrożony upadłością, gdy sprawozdania finansowe są składane z dużym opóźnieniem. Jeżeli mamy obserwację B , to prawdopodobieństwo $P(A|B)$ jest otrzymane z reguły Bayesa: tj z pomnożenia $P(A)$ przez $P(B|A)/P(B)$. Zbiór uczący służy do wyliczenia prawdopodobieństw warunkowych.

Sieci Bayesa mają kilka zalet:

1. W przeciwieństwie do innych technik regresji nie wymagają określenia rozkładów zmiennych użytych w modelu.
2. W sieciach można łatwo konstruować relacje między zmiennymi włączając pośrednich mediatorów i uzyskując efekty interakcji. Te relacje są przekładane na model użytkownika. Ta cecha jest ważna w dziedzinie prognozowania zagrożenia upadłością gdyż między zmiennymi objaśniającymi często występują wzajemne relacje. Dla przykładu wskaźniki finansowe i opinie audytorów mogą być użyte jako zmienne objaśniające. Opinie audytorów mogą być wyjaśniane przez finansowe informacje.
3. Sieci Bayesa nie wymagają kompletnych informacji z poszczególnych obserwacji. W obserwacjach mogą nie występować pewne zmienne. To jest bardzo ważne w analizie firm upadających, ponieważ te firmy mają skłonność do ukrywania informacji.
4. Sieci BN stanowią modele dynamiczne i interaktywne. Można do nich łatwo wprowadzić nową informację w procesie uczenia. Również można wprowadzić łatwo subiektywną wiedzę ekspertów.
5. Raporty dla użytkowników są łatwe do zrozumienia gdyż reprezentowane graficznie przez skierowany acykliczny graf. W modelu opracowanym przez L.Sun, P.P. Shenoy punkt odcięcia (cut off point) wynosi 0,5. To znaczy, że jeżeli prawdopodobieństwo stanu za-

grożenia upadłością jest większe niż 0,5, to dane przedsiębiorstwo uznane jest za bankruta.

Najbardziej znanymi oprogramowaniami do klasyfikacji metodą sieci przekonania Bayesa są: NETICA APPLICATION firmy Norsys Software Corp.¹⁹ oraz oprogramowanie HUGIN firmy duńskiej Hugin Expert A/S²⁰.

W środowisku bazy danych ORACLE 9i Release 2 istnieją mechanizmy umożliwiające przeprowadzenie procesu drażenia danych. Do klasyfikacji służą: adaptacyjna sieć Bayesa (Adaptive Bayes Network), naiwny klasyfikator Bayesowski (Naive Bayes Network) oraz poszukiwacz modelu (Model Seeker).²¹

Innym podejściem do maszynowego uczenia jest paradygmat indukcyjny. Do najbardziej znanych technik należą techniki stosujące algorytmy drzew decyzyjnych:

- CHAID (Chi-square automatic interaction detection),
- CART (Classification and Regression Trees),
- C4.5 oraz ID3.
- RPA (Recursive Partitioning algorithm).

Algorytm RP jest nieparametryczną metodą klasyfikacji, której istota polega na generowaniu drzewa klasyfikacyjnego. Metoda ta polega na rekurencyjnym podziale próby na dwie podpróby (lewą i prawą) na podstawie jednej ze zmiennych objaśniających wybranych do modelu. Należy tu podać jedną z cech tego podejścia, że nie ma tu konieczności wcześniejszej selekcji zmiennych. Algorytm wybiera zmienne z ustalonej z góry grupy zmiennych. Ta cecha jest szczególnie korzystna, jeżeli istnieje duża liczba zmiennych. Algorytm RP wybiera zmienne, które najlepiej klasyfikują obserwacje i optymalną wartość podziału zakresu wartości zmiennej. Kryterium optymalizacyjnym przy konstrukcji drzewa o optymalnej strukturze jest krzyżowe ryzyko podstawienia. Kryterium to pozwala na otrzymanie drzewa o odpowiedniej złożoności. Na podstawie zbudowanego drzewa przy pomocy algorytmu RP określony zostaje model typu z-score, czyli funkcja dyskryminacyjna określona na liniowej kombinacji wszystkich zmiennych wyodrębnionych w wierzchołkach drzewa. Współczynniki przed zmiennymi stanowią wartość podziału na grupę przedsiębiorstw zagrożonych i nie zagrożonych upadłością wg odpowiedniej zmiennej.

Zazwyczaj programy komputerowe bazujące na metodzie drzew decyzyjnych ilustrują dokładnie drzewo graficznie. W węzłach podawane są następujące informacje:

- numer kolejny wierzchołka,
- wartość rozgraniczającą zmienną objaśnianą,
- dane statystyczne,
- % złej klasyfikacji.

¹⁹ WWW.norsys.com/netica.html

²⁰ www.hugin.com/Products_Services/

²¹ Kościów S., Pondel M., Kotwica A. (2003) Zastosowanie technologii drażenia danych w systemach klasy CRM w oparciu o środowisko ORACLE 9i. Prace Naukowe Akademii Ekonomicznej we Wrocławiu, nr 975, Pozyskiwanie wiedzy i zarządzanie wiedzą. 228-238.

¹⁸ Sun Lin, Shenoy P. (2003) Using Bayesian Networks for bankruptcy prediction in stressed firms. School of Business Working Paper no.295 University of Kansas, www.aaahq.org/infosys/index.htm

Dostępne oprogramowanie DTREG²² obejmuje problematykę generowania drzew klasyfikacyjnych i regresyjnych. Można przeprowadzić analizę zbioru drzew decyzyjnych „lasów”. Na podstawie algorytmu J. H. Friedmana program buduje model regresyjny. Technika TreeBoost pozwala na optymalizację dokładności klasyfikacji.

T. K. Sung, N. Chang, G. Lee²³ dokonali porównania technik typu „data mining” stosowanych do prognozowania stanu upadłości przedsiębiorstw. Technika drzew decyzyjnych posiada następujące zalety:

- zdolność do generowania zrozumiałych reguł,
- zdolność do operowania zarówno ciągłymi jak i dyskretnymi zmiennymi,
- wskazanie które obszary są ważniejsze dla prognozy i klasyfikacji,
- łatwość obliczeń i generowania wyników w graficznej postaci.

Inni autorzy²⁴ dokonują porównania oprogramowania See 5, PolyAnalyst i Cubist, w których stosuje się techniki drzew decyzyjnych.

Inną techniką stosowaną wg paradygmatu indukcyjnego jest metoda wyodrębniania reguł decyzyjnych. Metoda ta stanowi podstawę tworzenia tzw. systemów ekspertowych. W metodzie generowane są reguły typu:

```
IF bieżące aktywa/bieżące zobowiązania < 1,533 ORAZ
  IF kapitał własny/aktywa ogółem < 0,092 ORAZ
    IF zysk zatrzymany/aktywa ogółem < 0,064 ORAZ
      IF zysk zatrzymany/ aktywa ogółem < -0,942
        THEN firma nie będzie zagrożona
      ELSE firma będzie zagrożona
    ELSE firma nie będzie zagrożona
  ELSE firma nie będzie zagrożona
ELSE firma nie będzie zagrożona.
```

Opracowano wiele algorytmów maszynowego uczenia do generowania reguł decyzyjnych. Algorytm „Explore” w zastosowaniu do oceny ryzyka kredytowania przedstawiają J. Stefanowski, S. Wilk²⁵ z Politechniki Poznańskiej.

Dostępny program jest XpertRule Knowledge Builder²⁶. Istnieje wiele różnych algorytmów z wy-

korzystaniem teorii zbiorów rozmytych. M. Spanos, G. Dounias oraz N. Matsatsinis wykorzystali do obliczeń oprogramowanie firmy MIT GmbH DataEngine oraz WINROSA²⁷. W podejściu wykorzystującym teorię zbiorów rozmytych tworzone są reguły ze zmiennymi lingwistycznymi²⁸:

```
IF FC jest dobra, GM jest zły oraz CP jest wysoka
THEN SCORE jest średnie.
```

FC- sytuacja finansowa, GM- stan zarządzania, CP- charakterystyka perspektyw
SCORE –ocena punktowa stanu zagrożenia.

M. Wasiak z Akademii Ekonomicznej we Wrocławiu²⁹ opracował model wnioskowania rozmytego z bazy analiz fundamentalnych spółek giełdowych znajdujących się na giełdzie warszawskiej (1997-2001). Autor określił pięć kategorii zagrożenia upadłością: 1 (kondycja bardzo słaba) do 5 (kondycja bardzo dobra).

Innym podejściem do budowy systemów opartych o regałową bazę wiedzy jest wykorzystanie teorii zbiorów przybliżonych³⁰.

Reguła decyzyjna ma postać:

```
IF koniunkcja atrybutów warunkowych
THEN dysjunkcja atrybutów decyzyjnych.
```

Atrybuty warunkowe i decyzyjne są podane w tablicy decyzyjnej.

T.E. Mekce z East Tennessee State University³¹ przeprowadził badania na próbie 100 amerykańskich przedsiębiorstw sklasyfikowanych do grupy zagrożonych upadłością i 100 przedsiębiorstw nie zagrożonych. Jako zmienne przyjął osiem wskaźników finansowych dla których przyporządkował wartości rozmyte: 1-wysoka wartość wskaźnika, 2-średnia, 3-niska. Do analizy danych zastosował oprogramowanie z Politechniki Poznańskiej: RoughDas Rough Set Based Data Analysis System Version 2.0. Oprogramowanie ROSE z Politechniki Poznańskiej wykorzystali hiszpańscy naukowcy do prognozowania stanu

²² www.dtreg.com/

²³ Sung T.K., Chang N., Lee G. (1999) Dynamics of modeling in data mining: interpretive approach to bankruptcy prediction. Journal of Management Information Systems, vol.16, no.1, pp.63-85.

²⁴ Norman E., Tan C., Kumar K., Torbey V., Drew D. Return prediction models and software reviews. www.erickas.net/write.php

²⁵ Stefanowski J., Wilk S. (2001) Evaluating business credit risk by means of approach- integrating decision rule and case-bases learning. International Journal of Intelligent Systems in Accounting, Finance & Management, 10, 97-114 www.attar.com/pages/info_kb.htm

²⁶ Spanos M., Dounias G., Matsatsinis N. A fuzzy knowledge-based decision aiding method for the assessment of financial risk: the case of corporate bankruptcy prediction. www.erudit.de/erudit/events/esit99/12621_p.pdf

²⁶ www.attar.com/pages/info_kb.htm

²⁷ Spanos M., Dounias G., Matsatsinis N. A fuzzy knowledge-based decision aiding method for the assessment of financial risk: the case of corporate bankruptcy prediction. www.erudit.de/erudit/events/esit99/12621_p.pdf

²⁸ Khan A.H. (2002) Can Banks learn to be rational. www.e.u-tokyo.ac.jp/cirje/research/dp/2002/2002cf151.pdf

²⁹ Wasiak M. (2002) Rozmyty system wnioskujący, oceniający kondycję finansową przedsiębiorstw. Pozyskiwanie wiedzy z baz danych. Wyd. Akademii Ekonomicznej nr 931, Wrocław.

³⁰ Mrózek A., Płonka L. (1999) Analiza danych metodą zbiorów przybliżonych. Akademicka Oficyna Wydawnicza PLJ Warszawa.

³¹ Mekce T.E. (2000) Developing a bankruptcy prediction model via rough sets theory. International Journal of Intelligent Systems in Accounting Finance & Management 9, 159-173.

instytucji ubezpieczeniowych³². Tay F., Shen L.³³ przedstawili stan zastosowań modelu opartego na teorii zbiorów przybliżonych w prognozach ekonomicznych i finansowych. Przedstawili analizę oprogramowania opracowanego w kilku ośrodkach akademickich (LERS, LEM2, RSES, TRANCE, DataLogic ProbRough, RoughDas, Profit).

Paradygmat ewolucyjny uczenia próbuje zastosować darwinowskie idee ewolucji i drogi doboru naturalnego do różnorodności zadań uczenia się. Pojęcie klasyfikatora w tym kontekście jest równoważne regule IF... THEN... Od klasyfikatora pomyślnego oczekuje się, by szybko osiągnął stan celu, omijając obszary przestrzeni przeszukiwań, które nie wydają się być obiecujące. Operatory genetyczne (mutacja, krosowanie) są stosowane do klasyfikatorów by generować nowe. Pomiędzy konkurującymi klasyfikatorami te, które osiągają wysokie rezultaty spowodowane dopasowaniem „przetrwają” i „rozmnażają się” by produkować następną generację. Poprzez ten proces, przestrzeń przeszukiwań jest zapełniona przez pasujące hipotezy, a szansa naprowadzenia na pojęcie docelowe wzrasta. Ta cecha czyni programowanie ewolucyjne procesem bardziej wyrafinowanym, niż losowe przeszukiwanie kierujące się tylko zasadą „zachowaj najlepszy”. Główne zastosowania algorytmów genetycznych dotyczą prognozowania cen akcji na giełdach.³⁴

Paradygmat konekcyjny reprezentują sztuczne sieci neuronowe. Uczenie się jest uzyskiwane tu poprzez modyfikację parametrów wagowych połączeń między neuronami. Dla ustalonego zbioru przykładów trenujących, zadaniem algorytmu uczenia się jest „zbieganie się” do zbioru wag, które aproksymują (nieznana) funkcję na tych przykładach. Wsteczna propagacja błędu jest strategią najbardziej szeroko używaną do tego celu. Wadą tego podejścia jest brak zdolności rozumowania o wyjściu w sposób, który jest efektywnie zrozumiały. Przegląd zastosowań sieci neuronowych w prognozie ryzyka kredytowego przedstawił A. F. Atija³⁵. Analizę dokładności prognozy przy użyciu sieci neuronowych przedstawił M. Anandarajan, P. Lee, A. Anandarajan³⁶.

³² Segovia-Vargas, Fana J.A., Heras-Martines A., Vilar-Zanon J.L., Sanchis-Arellano A. Using rough sets to predict insolvency of Spanish non-live insurance companies. www.ucm.es/BUCEM/cee/doc/03-002/03002.pdf

³³ Tay F.E., Shen L. (2002) Economic and financial prediction using rough sets model. *European Journal of Operational Research* 141 641-659.

³⁴ Shin K., Kim K., Han I. Financial Data Mining using genetic algorithms technique: application to KOSPI 200, http://afis.kaist.ac.kr/download/inner_con053.pdf

³⁵ Atija A.F. (2001) Bankruptcy prediction for credit risk using neural networks: a survey and new results. *IEEE Transactions on Neural Networks*, vol.12, no.4, 929-935.

³⁶ Anandarajan M., Lee P, Anandarajan A. (2001) Bankruptcy prediction of financially stressed firms: an examination of the predictive accuracy of artificial neural networks. *International Journal of Intelligent Systems in Accounting, Finance & Management*, 10, 69-81.

Przesłanką syntaktycznego (strukturalnego) rozpoznawania wzorca jest stwierdzenie, że strukturalne relacje między składnikami wzorców są decydujące dla celów klasyfikacji i opisu. Metodologia syntaktycznego uczenia się jest realizacją modelu lingwistycznego, który został zaproponowany by uzupełnić klasyfikację statystyczną; wynika to z faktu, że model statystyczny był krytykowany za to, iż jest tylko metodą przydziału klas bez oferowania realnej wiedzy o klasie (np. jej opisu).

Gramatyki były najbardziej szeroko używanymi środkami reprezentacji dla celów rozpoznawania syntaktycznego. Wzorzec jest rozpoznany jako członek klasy, jeżeli może być wyprowadzony używając reguł produkcji dla gramatyki odpowiadającej tej klasie. Wyprowadzenie jest wykonywane przez analizę składniową (parsing). Możliwe jest uzyskanie pewnej elastyczności takich systemów poprzez wprowadzenie reguł, które są etykietowane ich prawdopodobieństwami (gramatyki stochastyczne).

Podejście to wykorzystali do prognozowania zagrożenia upadłością przedsiębiorstw A. Brabazon, M. O'Neill³⁷. Przedstawili oni kilka gramatyk. Jedna z nich ma następujący układ:

```
<lc> ::= ouput = <expr> ;
<expr> ::= (<expr>) + (<expr>)
           | <coeff> * ( <var> / <var> )
< var > ::= Sales [ Net Income [ Gross Profit
           | EBIT | EBITDA | Total Assets
           | Total Current Assets | Total
           Liabilities
           | Total Current Liabilities | Total
           Long Term Debt
           | Cash From Operations | Free Cash Flow
,coeff. ::= ( <coeff> ) <op> ( <coeff> )
           | <float>
< op > ::= +
           | -
< float > ::= 20 | - 20 | 10 | -10 | 5 | 5 | 4 | - 4
           | 3 | - 3 | 2 | - 2 | 1 | - 1 |
           .1 | - .1
```

Zespół irlandzki przedstawił wyniki swoich badań w zakresie podejścia syntaktycznego, w których wykorzystano listę 22 wskaźników finansowych w gramatyce formalnej³⁸.

Paradygmat analityczny zawiera podejście dedukcyjne, które aby wydobyć opis pojęcia ogólnego opiera się na bogatej teorii dziedziny i tylko kilku przykładach (możliwie, że na jednym). Każdy przykład może być sko-

³⁷ Brabazon A., O'Neill (2004) Diagnosing corporate stability using grammatical evolution. *Int.J.Appl.Math.Comput.Sci.* Vol.14, No.3, 363-374.

³⁸ Brabazon A., Matthews R., O'Neill M., Ryan C. Grammatical Evolution and Corporate Failure Prediction. www.business.king.ac.uk/research/intbus/paper4.pdf

jarzony ze ścieżką wyvodu jego poprawności, zgodnie z wiedzą „wrodzoną”, która jest reprezentowana przez teorię dziedziny. Stosuje się tu uczenie w oparciu o wyjaśnianie (explanation-based learning).

Metodologią maszynowego uczenia jest Case Based Reasoning. Stosuje się tu trzy fazy:

- wprowadzenie początkowej wiedzy lub przykład problemu,
- zdefiniowanie sugerowanych rozwiązań problemu,
- załadowanie pamięci studiami przypadków celem uruchomienia mechanizmu wnioskowania.

C.Park, I. Han³⁹ zaprezentowali metodę CBR z wykorzystaniem algorytmu k-NN do określenia podobieństwa między przypadkiem źródłowym S oraz przypadkiem rozpatrywanym. Do określenia wag poszczególnych cech obiektu użyto metody AHP (Analytic Hierarchy Process⁴⁰) oraz oprogramowania Expert Choice⁴¹. Autorzy przeprowadzili badania na 2144 przypadkach z Industrial Bank of Korea. Do oceny wag przyjętych cech przedsiębiorstw wykorzystano 21 ekspertów.

Trzecią kategorią modeli są modele opracowane na gruncie teorii wielokryterialnego wspomaganie decyzji. Rozważane są tu cztery problematyki wzorcowe⁴²:

- problematyka wyboru: wspomaganie wyboru „najlepszego” wariantu lub wspomaganie opracowania procedury selekcji,
- problematyka sortowania: wspomaganie sortowania wariantów według norm lub wspomaganie opracowania procedury przydziału,
- problematyka porządkowania: wspomaganie porządkowania wariantów według malejącej preferencji lub wspomaganie opracowania procedury klasyfikacji,
- problematyka opisu: wspomaganie opisu wariantów i ich konsekwencji w sposób systematyczny i sformalizowany lub wspomaganie opracowania procedury poznawczej.

Problematyka sortowania polega na postawieniu problemu przydziału wariantów ze zbioru A do określonych kategorii, to znaczy na skierowaniu badań na poszukiwanie podzbioru zbioru A na kategorie według norm odnoszących się do immanentnych wartości rozpatrywanych wariantów. Uwzględnia się przy tym charakter rewidowany i przejściowy zbioru A. Problematyka ta przewiduje formę zalecenia lub zwykłego udziału, który ma na celu:

- zalecenie akceptacji pewnych wariantów lub ich odrzucenia, przy czym pozostałe mają być przed-

miotem dalszej analizy z uwzględnieniem koncepcji kategorii, lub

- zaproponowanie przyjęcia ogólnej metodyki opartej na procedurze przydziału do kategorii wszystkich wariantów, nadającej się do wielokrotnego i automatycznego zastosowania.

Stan badań dotyczących zastosowania do prognozowania stanu upadłości metod wypracowanych w obszarze teorii wielokryterialnego wspomaganie decyzji przedstawili w swojej książce C. Zopounidis i A. I. Dimitras⁴³.

Najbardziej reprezentatywną metodą sortowania jest metoda ELECTRE TRI. Metoda ta wywodzi się ze zbioru metod ELECTRE, których koncepcje przedstawili uczeni francuscy z University of Paris Dauphine. Procedurę wnioskowania z przykładów dla tej metody przedstawili A. Ngo i V. Rousseau⁴⁴. ELECTRE TRI jest wielokryterialną metodą sortowania stosowaną do przydziału wariantów do zdefiniowanych uprzednio i uporządkowanych kategorii. Przydział wariantu „a” wynika z porównania „a” z profilami definiującymi granice kategorii. Każdy profil „r” charakteryzowany jest przez zestaw wartości kryterium „g”, preferowaną wartość „p”, wartość neutralną „q”. Praktycznie wartość neutralna dla danego kryterium oznacza maksymalną akceptowalną różnicę między wartością kryterium określoną dla profilu „r” a wartością kryterium określoną dla danego wariantu. Wartość preferowana jest rozumiana jako maksymalna akceptowana wartość między wartością profilu „r” a wartością kryterium dla danego wariantu „a”, w przypadku kiedy rozważamy pewną różnicę w preferencji między nimi.

W metodzie ELECTRE TRI celem przeprowadzenia sortowania obliczony jest indeks zgodności $c(a, r)$ wyrażający siłę stwierdzenia: „wariant „a” jest co najmniej tak dobry jak profil „r” wg danego kryterium”. W celu porównania dwóch wariantów na podstawie więcej niż jednego kryterium obliczany jest globalny indeks zgodności $C(a, r)$. W tej metodzie wyliczany jest również indeks niezgodności $D(a, r)$ wyrażający siłę zaprzeczenia stwierdzenia: „ wariant „a” jest co najmniej tak dobry jak profil „r” zgodnie z kryterium „g””. Ostatecznie wyliczany jest indeks wiarygodności $s(a, r)$, służący do określenia relacji przewyższania S.

Wariant „a” przewyższa profil „r” jeżeli indeks wiarygodności $s(a, r)$ jest większy od wartości „lambda” stanowiącej tzw. poziom graniczny (cut level). Wartość ta jest ustalona w przedziale domkniętym od 0,5 do 1.

Ponadto metoda zawiera dwie procedury: pesymistyczną i optymistyczną pozwalające na analizę wrażli-

³⁹ Park C.S., Han I. (2002) A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction. *Expert Systems with Applications*, 2-10.

⁴⁰ Saaty T. (2001) *Decision Making for Leaders. The analytic hierarchy process for decisions in complex World*. RWS Publications.

⁴¹ Forman E., Selly M. (2001) *Decision by objectives*. World Scientific.

⁴² Roy Bernard (1990) *Wielokryterialne wspomaganie decyzji*. WNT Warszawa.

⁴³ Zopounidis C., Dimitras I. (1998) *Multicriteria decision aid methods for the prediction of business failure*. Kluwer Academic Publishers.

⁴⁴ Ngo An, Rousseau Vincent (2002) using assignment examples to infer category limits for the ELECTRE TRI method, *Journal of Multi-Criteria Decision Analysis* 11, 29-43.

wości. Wg obydwu procedur dokonywany jest przydział analizowanych wariantów do poszczególnych kategorii.

Dostępne jest oprogramowanie metody ELECTRE TRI Software 2.0 opracowane w Politechnice Poznańskiej i Université Paris Dauphine przez zespół V. Mousseau, R. Słowiński, P. Zielniewicz⁴⁵. Innym oprogramowaniem jest portugalski system IRIS (Inferencie for multicriteria sorting problems ver.2)⁴⁶ opracowany w Institute of Systems Engineering and Computers INESC Coimbra, Portugal.

W analizie stanu upadłości przedsiębiorstw metodę ELECTRE TRI zastosowali Dimitras i Zopounidis przeprowadzając badania dla Greek Bank. J. Khalil, J. Martel, P. Juras⁴⁷ przeprowadzili badania dotyczące określenia ryzyka kredytowego w Kanadzie.

W metodzie można zdefiniować kilka stanów zagrożenia upadłością. Grecy zdefiniowali trzy: C1 - grupa wysokiego ryzyka, C2 - grupa przedsiębiorstw niepewnych, C3 - grupa niskiego ryzyka. Kanadyjczycy określili pięć stanów: C1 (ryzyka), C2 (wysokiego ryzyka), C3 (ryzyka akceptowanego), C4 (normalnego ryzyka), C5 (niskiego ryzyka).

Czwarta grupa modeli dotyczy określenia wiarygodności kredytowej. Rating obejmuje wyznaczenie syntetycznego wskaźnika określającego stopień zaufania do danego przedsiębiorstwa, banku, państwa. Kryteria oceny podmiotów gospodarczych uwzględniane przy opracowywaniu ratingów są zróżnicowane ze względu na: typ działalności, przynależność branżową, cechy lokalne, warunki prawne, oraz uwarunkowania konkurencyjne. W metodologii agencji ratingowej Moody's KMV określona zostaje miara prawdopodobieństwa utraty wypłacalności EDF (Expected Default Frequency)⁴⁸. Opracowano również zalecenia standardowe Bazylei II.

Altman⁴⁹ przedstawił analizę zastosowania modeli od analizy dyskryminacyjnej do standardowych wskaźników KMV i BASEL II.

Podsumowanie

Z przedstawionej w niniejszej pracy klasyfikacji metod wynika, że modelowanie problematyki prognozowania stanu zagrożenia upadłością odbywa się w oparciu

o różne teorie. Najwięcej modeli opracowano z wykorzystaniem statystyki matematycznej, maszynowego uczenia oraz wielokryterialnego wspomaganie decyzji. Dlatego do badań należy zastosować te trzy podejścia.

W modelach statystycznych główną wagę przywiązuje się do klasyfikacji. Modele te cechuje duża dokładność klasyfikacji. Istnieje tu skłonność do redukcji zmiennych objaśniających, gdyż celem nie jest wyjaśnianie wpływu zmiennych objaśniających na zmienną objaśnianą, ale dokładność klasyfikacji.

W modelach wykorzystujących techniki maszynowego uczenia podstawowym zagadnieniem jest interpretacja na podstawie analizy danych. Modele te ukierunkowane są na generowanie reguł objaśniających. Graficzna interpretacja np. drzew decyzyjnych wspomaga proces interpretacji. Zastosowanie teorii zbiorów przybliżonych umożliwia analizę w przypadku niejednoznacznego przyporządkowania obiektów do różnych klas.

Zastosowanie teorii zbiorów rozmytych umożliwia definiowanie cech obiektów w kategoriach zmiennych lingwistycznych, co umożliwia wprowadzanie dodatkowych interpretacji ekspertów oraz modelowanie ich ocen.

Problematyka sortowania wprowadzona na grunt wielokryterialnego wspomaganie podejmowania decyzji stanowi odrębne podejście wymagające od ekspertów zdefiniowania kategorii klasyfikacyjnych. To podejście wymaga szerokiej analizy każdej z cech obiektu z uwagi na określenie wzorca reprezentującego klasę. Takie podejście wymaga stosowania benchmarkingu w zakresie analizy różnych stanów zagrożenia upadłością przedsiębiorstw.

Wykorzystanie powyższych metod do konstrukcji regionalnego systemu wczesnego ostrzegania przed upadłością pozwoli na integrację różnych badań przeprowadzonych na poziomie makroekonomicznym, mezoekonomicznym i mikroekonomicznym. Dla przedsiębiorców i innych interesariuszy ważnym będzie analiza przyczynowo-skutkowa, którą można będzie przeprowadzić o wyniki z modeli interpretacyjnych.

Odrębnym zagadnieniem jest porównanie dokładności prognozowania wg danego modelu.

⁴⁵ Rousseau V., Słowiński R., Zielniewicz P. (2000) A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. *Computers & Operations Research* 27, 757-777.

⁴⁶ Dias L., Mousseau V. (2003) IRIS User Manual Documents of INESC Coimbra No.1/2003, www.lamsade.dauphine.fr/~mousseau/dias3docl02.pdf

⁴⁷ Khalil J., Martel J.M., Juras P. (1999) A multicriterion system for credit risk rating, Université Laval, Document de Travail 1999-014, <http://archmede.bibl.ulaval.ca/di/files/222/1-3-222-20040428-1.pdf>

⁴⁸ Jajuga K., Krysiak Z. (red.) (2004) Ryzyko kredytowe wiarygodności hipotecznych, Związek Banków Polskich, Warszawa.

⁴⁹ Altman E. (2002) Corporate distress prediction models in a turbulent economic and BASEL II environment. www.pages.stern.nyu.edu/~ealtman/Corp-Distress.pdf