

Józef Pociecha

Uniwersytet Ekonomiczny w Krakowie

WYBRANE METODY KLASYFIKACYJNE ORAZ ICH EFEKTYWNOŚĆ W PROGNOZOWANIU UPADŁOŚCI FIRM

Wprowadzenie

„Tworzenie pojęć oraz reguł pozwalających klasyfikować obiekty na klasy podobieństwa odpowiadające tym pojęciom towarzyszy rozwojowi każdej dyscypliny naukowej. Proces badania, czy dowolny obiekt należy do danej klasy, nazywamy rozpoznawaniem obiektu” (Kolonko, 1982, s. 7). Takim stwierdzeniem rozpoczyna się podstawowe dzieło profesora Józefa Kolonki, wprowadzające czytelnika w szeroką, a w tym czasie mało znaną problematykę klasyfikacji zadań, metod klasyfikacji oraz zagadnień redukcji opisu. Ta niewątpliwie nowatorska w tym czasie praca dała impuls do intensywnego rozwoju metodologii klasyfikacji danych i stała się inspiracją do publikowania dalszych prac z zakresu taksonomii, metod dyskryminacyjnych oraz metod rozpoznawania obrazów. Praca prof. Kolonki wskazywała jednocześnie na szerokie możliwości wykorzystania proponowanej metodologii badawczej do wielorakich analiz ekonomicznych, m.in. do normowania kosztów, konstrukcji grupowych modeli konsumpcji, analizy danych przekrojowo-czasowych oraz porównywania i porządkowania obiektów wielocechowych. Mimo upływu lat od czasu jej opublikowania ciągle inspiruje do rozwoju metod klasyfikacji danych oraz ich zastosowań w badaniach społeczno-ekonomicznych.

Artykuł ten nawiązuje do problematyki metodologicznej pracy prof. Kolonki oraz jest kolejnym przykładem możliwości zastosowania metod klasyfikacji danych we współczesnych analizach ekonomicznych. Przedmiotem artykułu jest przedstawienie najpopularniejszych metod prognozowania bankructwa, będących w istocie statystycznymi lub iteracyjnymi metodami klasyfikacji danych, oraz ocena ich efektywności jako narzędzia oceny ryzyka upadłości firm.

Według przeprowadzonych badań (Aziz, Dar, 2004) jako narzędzia prognozowania upadłości firm najczęściej stosowane są modele dyskryminacyjne (30,3% publikowanych prac), na drugim miejscu pod względem częstości zastosowań znajdują się modele logitowe (21,3%), na trzecim miejscu sieci neuron-

we (9,0%) a na czwartym miejscu – drzewa klasyfikacyjne (5,6%). Te też narzędzia zostaną pokrótce scharakteryzowane, a w dalszej kolejności zostaną przytoczone ich pierwotne zastosowania znane z literatury światowej oraz wybrane zastosowania dla celów predykcji bankructwa firm we współczesnej gospodarce polskiej. W końcowej części pracy przedstawiono ocenę efektywności przedstawionych metod jako narzędzi prognozowania bankructwa. Ich efektywność została zmierzona zdolnością do poprawnej klasyfikacji firm do zbioru bankrutów i niebankrutów na zbiorze uczącym oraz zbiorze testowym. W zakończeniu sformułowano pewne ogólne warunki określające przydatność prezentowanych metod do predykcji bankructwa firm.

1. Liniowa funkcja dyskryminacyjna

Liniowa funkcja dyskryminacyjna, jako narzędzie klasyfikacji danych, została sformułowana w 1936 r. przez R.A. Fishera (1936) i zastosowana w badaniach przyrodniczych. Poniżej zostanie przedstawiona jej zasadnicza idea. W ujęciu klasycznym* przyjmujemy, że mamy dwie populacje. Załóżmy że $X_1 = [x_1, x_2, \dots, x_{n_1}]$ jest wektorem n_1 obserwacji z populacji 1 a $X_2 = [x_{n_1+1}, x_{n_1+2}, \dots, x_{n_1+n_2}]$ jest wektorem n_2 obserwacji, pochodzącym z populacji 2. Są to wektory o wymiarach $p \times 1$, gdzie p jest liczbą zmiennych dyskryminujących. Fisher zaproponował liniową funkcję (dyskryminator) dla zaklasyfikowania wylosowanej obserwacji do jednej z dwóch wyróżnionych populacji. Metoda ta polega na znalezieniu liniowej transformacji oryginalnych zmiennych:

$$l(X) = a^t X \quad (1)$$

tak aby było możliwe ich maksymalne odseparowanie w dwóch populacjach. Zaproponował znalezienie wektora \hat{a} maksymalizującego funkcję separacji $|S(a)|$, gdzie:

$$S(a) = \frac{\bar{y}_1 - \bar{y}_2}{S_y}, \quad (2)$$

\bar{y}_1 oraz \bar{y}_2 są średnimi zmiennych transformowanych Y_1 z populacji 1 oraz Y_2 z populacji 2,

* Spośród wielu prac z tego zakresu w języku angielskim patrz np.: (Christensen, 1991; Giri, 1996 lub Rencher, 1998). Wprowadzenie do tej problematyki wraz z zastosowaniami do prognozowania bankructwa można znaleźć m.in. w pracach: (Pocięcha, 2006 lub Pocięcha, 2007).

$$S_y = \left[\frac{\sum_{i=1}^{n_1} (y_i - \bar{y}_1)^2 + \sum_{i=n_1+1}^{n_1+n_2} (y_i - \bar{y}_2)^2}{n_1 + n_2 - 2} \right]^{\frac{1}{2}} \quad (3)$$

oraz

$$y_i = a^t x_i, \quad i = 1, 2, \dots, n_1 + n_2. \quad (4)$$

$S(a)$ dane wzorem (2) mierzy różnicę pomiędzy przekształconymi średnimi $\bar{y}_1 - \bar{y}_2$ relatywnie do wielkości odchylenia standardowego z prób (3). Jeśli zmienne przekształcone y_1, y_2, \dots, y_{n_1} oraz $y_{n_1+1}, y_{n_1+2}, \dots, y_{n_1+n_2}$ byłyby całkowicie odseparowane, to $|\bar{y}_1 - \bar{y}_2|$ powinno być możliwie duże.

Wektor \hat{a} maksymalizujący separację $|S(a)|$ jest następujący:

$$\hat{a} = S^{-1}(\bar{x}_1 - \bar{x}_2) \quad (5)$$

gdzie:

$$S = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2} \quad (6)$$

$$S_1 = \frac{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)(x_i - \bar{x}_2)}{n_1 - 1} \quad (7)$$

$$S_2 = \frac{\sum_{i=n_1+1}^{n_1+n_2} (x_i - \bar{x}_1)(x_i - \bar{x}_2)}{n_2 - 1} \quad (8)$$

oraz \bar{x}_1, \bar{x}_2 są średnimi wektora 1 i 2.

Przyjmijmy że mamy obserwację x_0 . Opierając się na funkcji dyskryminacyjnej (1), możemy alokować tę obserwację do odpowiedniej populacji na podstawie następujących reguł klasyfikacyjnych:

– alokuj x_0 do populacji 1, jeśli:

$$\hat{y}_0 = (\bar{x}_1 - \bar{x}_2)^t S^{-1} x_0 \geq \frac{1}{2} (\bar{x}_1 - \bar{x}_2)^t S^{-1} (\bar{x}_1 + \bar{x}_2) \quad (9)$$

– alokuj x_0 do populacji 2, jeśli:

$$\hat{y}_0 = (\bar{x}_1 - \bar{x}_2)' S^{-1} x_0 < \frac{1}{2} (\bar{x}_1 - \bar{x}_2)' S^{-1} (\bar{x}_1 + \bar{x}_2) \quad (10)$$

Innymi słowy, jeśli \hat{y}_0 znajduje się po prawej stronie $\frac{\bar{y}_1 + \bar{y}_2}{2}$ (bliżej do \bar{y}_1), zaklasyfikuj x_0 do populacji 1 i odwrotnie.

Analiza dyskryminacyjna jest od przeszło 40 lat wykorzystywana do prognozowania upadłości firm.

2. Model logitowy

Modele logitowe należą do klasycznych modeli klasyfikacji binarnej, tj. takich, w których zmienna objaśniana przyjmuje tylko dwie wartości. Przy analizie zagrożenia upadłością przedsiębiorstwa badamy wpływ zmiennych objaśniających, które mogą być cechami jakościowymi lub ilościowymi, na zmienną objaśnianą o charakterze jakościowym, przy założeniu logistycznej postaci analitycznej natężenia poszczególnych zmiennych objaśniających*. Jeśli zmienna objaśniana ma charakter zero-jedynkowy (przedsiębiorstwo o dobrej kondycji finansowej, przedsiębiorstwo o złej kondycji), to mamy do czynienia z modelem dychotomicznym. Gdy zmienna objaśniana jest cechą jakościową wielowariantową (bankrut, średnia kondycja finansowa, dobra kondycja finansowa), to mamy do czynienia z modelem wielomianowym uporządkowanym.

Wartości zmiennej objaśnianej wskazują na wystąpienie lub brak wystąpienia pewnego zdarzenia, które chcemy prognozować. Regresja logistyczna pozwala na obliczanie prawdopodobieństwa tego zdarzenia jako prawdopodobieństwa sukcesu według wzoru:

$$P(x) = \frac{e^x}{1+e^x} \quad (11)$$

lub

$$P(x) = \frac{1}{1+e^{-x}} \quad (12)$$

czyli funkcji logistycznej podającej prawdopodobieństwo bankructwa firmy.

Model regresji logistycznej należy do grupy uogólnionych modeli liniowych (GLM), w którym wykorzystano formułę logitu jako funkcji wiążącej (McCulloch, Searle, Neuhaus, 2009). Logitem nazywamy funkcję przekształcającą prawdopodobieństwo na logarytm ilorazu szans, czyli:

* Na temat modelu logitowego oraz możliwości jego zastosowania patrz np.: (Pocięcha, 2012).

$$L = \text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \ln(p) - \ln(1-p) \quad (13)$$

Dystrybuanta F prawdopodobieństwa sukcesu jest wzięta z rozkładu logistycznego, czyli:

$$p = F(\beta_0 + \beta_1 + \dots + \beta_k + \varepsilon) = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k + \varepsilon)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k + \varepsilon)} \quad (14)$$

Wartość funkcji odwrotnej do dystrybuanty F jest logitem (L) określonym wzorem (13). Logit jest więc logarytmem ilorazu szans bankructwa i „niebankructwa” firmy. Jeżeli te szanse są jednakowe, czyli $p = 0,5$, to logit jest równy zero, dla $p > 0,5$ logit jest dodatni, a gdy $p < 0,5$ – jest on ujemny.

Po przekształceniu logitowym można przystąpić do badania zależności pomiędzy wartościami logitu a zmiennymi objaśniającymi, będącymi odpowiednimi wskaźnikami finansowymi, przyjmując najczęściej model liniowy o postaci:

$$L = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k + \varepsilon \quad (15)$$

Parametrów powyższego modelu nie można szacować klasyczną metodą najmniejszych kwadratów, gdyż nie jest spełnione założenie stałości wariancji dychotomicznej zmiennej objaśnianej. W tym przypadku można stosować uogólnioną metodę najmniejszych kwadratów, a obecnie, szczególnie przy małej liczbie obserwacji stosuje się metodę największej wiarygodności.

3. Sieć neuronowa

Sztuczna sieć neuronowa jest techniką informatyczną wzorowaną na strukturze i sposobie działania układów nerwowych organizmów żywych*. Sztuczny neuron jest modelem swojego rzeczywistego odpowiednika. Jego zasadniczym celem jest przetworzenie informacji wejściowej, dostarczanej w postaci wektora skończonej liczby sygnałów wejściowych x_1, \dots, x_n w wartość wyjściową y . Przyjmuje się, że zarówno wartości wejściowe neuronu, jak i wartość wyjściowa mają postać liczb rzeczywistych. Z każdym wejściem neuronu związany jest współczynnik nazywany wagą. Współczynniki wagowe neuronu są podstawowymi parametrami wpływającymi na sposób funkcjonowania sztucznej komórki nerwowej.

* Charakterystyka sieci neuronowych i ich zastosowania dla celów prognozowania bankructwa została przedstawiona m.in. w pracy: (Pociecha, 2010b).

Oprócz n wejść oraz jednego wyjścia (y), występuje także wartość stała dla każdego neuronu w_0 (wyraz wolny), niezależna od danych wejściowych, nazywana biasem. W najprostszym przypadku, tj. neuronu liniowego, przetwarzanie sygnałów wejściowych odbywa się na podstawie obliczenia sumy ważonej wejść o wagach odpowiednio w_i , czyli:

$$\bar{y} = w_0 + \sum_{i=1}^n w_i x_i = w_0 + \mathbf{w}^T \mathbf{x} \quad (16)$$

gdzie:

$\mathbf{x} = [x_i]$ – wektor $[n \times 1]$ sygnałów wejściowych,

$\mathbf{w} = [w_i]$ – wektor $[n \times 1]$ wag, które z jednej strony wyrażają stopień ważności informacji przekazywanej i -tym wejściem, a z drugiej strony stanowią swojego rodzaju pamięć neuronu, zapamiętują bowiem związki zachodzące pomiędzy sygnałami wejściowymi a sygnałem wyjściowym.

W pewnych zastosowaniach przyjmowane są inne formuły agregacji, np. w postaci kwadratu odległości Euklidesa pomiędzy wektorem wejściowym \mathbf{x} a wektorem wag \mathbf{w} ; wtedy takie neurony są nazywane neuronami radialnymi (Witkowska, 2002).

Drugim etapem przekształcenia otrzymanej wartości pobudzenia jest wyznaczenie sygnału (wartości) wyjściowej. Elementem odpowiedzialnym za wykonanie tej czynności jest funkcja aktywacji:

$$y = \varphi(\bar{y}) \quad (17)$$

Funkcja aktywacji może być funkcją liniową lub nieliniową.

Spośród znanych funkcji aktywacji najczęściej przyjmuje się funkcję logistyczną:

$$\varphi(\bar{y}) = \frac{1}{1 + \exp(-\beta \bar{y})} \quad (18)$$

lub tangens hiperboliczny:

$$\varphi(\bar{y}) = \frac{\exp(\beta \bar{y}) - \exp(-\beta \bar{y})}{\exp(\beta \bar{y}) + \exp(-\beta \bar{y})} \quad (19)$$

W celu stworzenia sieci neuronowej łączy się neurony w określony sposób. Zwykle neurony wchodzące w skład sieci tworzą warstwy, z których pierwsza nosi nazwę warstwy wejściowej, ostatnia – warstwy wyjściowej, zaś wszystkie znajdujące się pomiędzy nimi określane są jako warstwy ukryte. Wartości wejściowe sieci wprowadzane są na wejścia neuronów warstwy wejściowej. Następnie, poprzez istniejące połączenia, wartości wyjściowe neuronów jednej warstwy przekazywane są na wejścia elementów przetwarzających kolejnej war-

stwy. Wartości uzyskane na wyjściach neuronów ostatniej warstwy są wartościami wyjściowymi sieci.

Zwolennicy sieci neuronowych jako narzędzia prowadzenia badań ekonomicznych wskazują na ich przewagę w stosunku do klasycznych modeli regresji, modeli dyskryminacyjnych oraz klasycznych modeli tendencji rozwojowych. Z tego też względu sieci neuronowe zalecane są także jako narzędzie prognozowania bankructwa. Jednakże sposób funkcjonowania sieci neuronowej, gwarantujący prawidłowe rozwiązywanie postawionych przed nią problemów, uzależniony jest od dwóch podstawowych czynników (Lula, 1999):

- wartości współczynników wagowych neuronów składających się na sieć,
- struktury (topologii) sieci, która określana jest przez liczbę warstw, liczbę neuronów w poszczególnych warstwach, sposób połączeń neuronów oraz przyjęty model neuronu (sposób agregacji danych wejściowych, rodzaj zastosowanej funkcji aktywacji).

Ze względu na architekturę sieci neuronowych można wyróżnić trzy jej główne grupy. Pierwszą z nich stanowią sieci jednokierunkowe. Typowym przykładem sieci jednokierunkowej jest perceptron wielowarstwowy. Drugą grupę stanowią sieci rekurencyjne, w których dopuszcza się występowanie cykli. Dynamika tego typu sieci jest zdecydowanie bardziej skomplikowana niż w przypadku sieci jednokierunkowych. Typowym przykładem sieci rekurencyjnej jest sieć Hopfielda (Tadeusiewicz, 1993). Trzecią grupę sieci neuronowych stanowią sieci komórkowe. W tej grupie sieci łączone są w dowolny sposób, ale tylko w obrębie sąsiedzkich węzłów. Typowym przykładem tego typu sieci jest sieć SOM Kohonena.

Po podjęciu decyzji co do wyboru właściwej architektury sieci neuronowej i jej zbudowaniu, należy rozpocząć proces przygotowania sieci do prawidłowego jej działania, zwany uczeniem sieci. Sieć uczy się prawidłowo działać na podstawie prezentowanych jej przykładów realizacji badanych obiektów lub zjawisk. Opierając się na przedstawionych rzeczywistych przypadkach, sieć stara się odkryć i zapamiętać ogólne prawidłowości charakteryzujące te obiekty lub kierujące przebiegiem badanych zjawisk. Rozpoznanie reguły sztuczna sieć neuronowa przechowuje w postaci zakodowanej w wartościach współczynników wagowych neuronów.

Zbiór danych wykorzystywany w trakcie uczenia sieci nazywamy zbiorem uczącym. Proces uczenia sieci uwalnia nas od uciążliwego tworzenia i zapisywania algorytmu wymaganego dla przetwarzania danych wejściowych, tak aby uzyskać pożądany wynik końcowy. Nie odbywa się to jednak bez kosztów, gdyż ceną jest długotrwały i wymagający dużej mocy obliczeniowych proces uczenia. Co więcej, proces uczenia sieci jest zawsze procesem indeterministycznym,

czyli wynik uczenia nie jest nigdy całkowicie pewny. Najczęściej stosowanym algorytmem uczenia z nauczycielem jest algorytm wstecznej propagacji błędu oparty na regule *delta* (Korbicz, Obuchowicz, Uciński, 1994).

W przypadku uczenia bez nauczyciela zbiór uczący zawiera tylko wartości zmiennych wejściowych. Uczenie polega na cyklicznym prezentowaniu danych uczących i na stopniowej, systematycznej modyfikacji wag, prowadzącej w efekcie do wytworzenia w sieci pewnej wiedzy o ogólnych cechach i właściwościach zbiorowości sygnałów wejściowych. Sieci uczone w trybie bez nauczyciela stosowane są do rozwiązywania zadań klasyfikacji bezwzorcowej, mającej na celu rozpoznanie struktury analizowanego zbioru obiektów lub identyfikacji jednorodnych fragmentów szeregów czasowych. Podstawowym algorytmem treningu sieci w trybie bez nauczyciela jest reguła Hebba (Ossowski, 1996).

4. Drzewo klasyfikacyjne

Drzewo klasyfikacyjne jest jednym z obrazów podziału rekurencyjnego badanego zbioru. Polega on na stopniowym podziale wielowymiarowej przestrzeni cech na rozłączne podzbiory, aż do uzyskania ich homogeniczności ze względu na wyróżnioną cechę (y). Następnie w każdym z uzyskanych segmentów budowany jest lokalny model tej zmiennej. Graficzną prezentacją metody podziału rekurencyjnego jest drzewo decyzyjne. Jeśli takie drzewo odnosi się do cechy y będącej cechą nominalną, to reprezentujące ją drzewo nazywane jest drzewem klasyfikacyjnym, a jeśli jest to zmienna ciągła, to takie drzewo nazywamy drzewem regresyjnym (Por. Gatnar, 2001, s. 26). W przypadku prognozowania bankructwa nasza cecha y jest cechą binarną (bankrut, niebankrut), dlatego też mówimy o drzewach klasyfikacyjnych jako narzędziu bankructwa.

Przedmiotem podziału rekurencyjnego (Por. Gatnar, 2001, s. 13) jest N -elementowy zbiór obiektów, w naszym przypadku firm, scharakteryzowanych przez wektor $M+1$ cech, będących wskaźnikami ich kondycji finansowej (płynność, zadłużenie, efektywność działania, rentowność): $[\mathbf{x}, y]$, gdzie:

$$\mathbf{X} = [x_1, x_2, \dots, x_M]. \quad (20)$$

Wielowymiarowe obserwacje wskaźników finansowych w analizowanym zbiorze firm można zapisać w postaci macierzy:

$$[\mathbf{x}_n, y_n]_{N \times M+1} = \begin{bmatrix} x_{11} & \dots & x_{1M} & y_1 \\ x_{21} & \dots & x_{2M} & y_2 \\ \dots & \dots & \dots & \dots \\ x_{N1} & \dots & x_{NM} & y_N \end{bmatrix} \quad (21)$$

Zgodnie z tradycyjnymi określeniami zmienne x_1, x_2, \dots, x_M są nazywane predyktorami, a zmienna y jest zmienną objaśnianą. Jest ona zmienną zero-jedynkową (0 – niebankrut, 1 – bankrut). Dysponując danymi z macierzy (21), należy znaleźć taką relację pomiędzy zmienną y a zmiennymi x_1, x_2, \dots, x_M , aby na podstawie znajomości wartości predyktorów można było określić wartość zmiennej y . Szuka się więc funkcji f , takiej że:

$$y = f(\mathbf{x}, \boldsymbol{\alpha}) + \varepsilon \quad (22)$$

W praktyce rozpatruje się jedynie model addytywny o postaci:

$$y = \alpha_0 + \sum_{k=1}^K \alpha_k g_k(\mathbf{x}, \boldsymbol{\beta}) \quad (23)$$

Stosując metodę rekurencyjnego podziału, otrzymujemy aproksymację modelu (23) w postaci funkcji:

$$y = a_0 + \sum_{k=1}^2 a_k I(\mathbf{x} \in R_k) \quad (24)$$

gdzie:

R_k (dla $k = 1, 2$) – rozłączne segmenty (bankrut, niebankrut) w wielowymiarowej przestrzeni cech,

a_k – parametry modelu.

Drzewo klasyfikacyjne jest graficzną reprezentacją modelu (24)*.

Model (24) nie jest tworzony globalnie, lecz poprzez złożenie modeli lokalnych, budowanych w każdym z K rozłącznych segmentów, na jakie dzielona jest wielowymiarowa przestrzeń zmiennych \mathbf{X}^m . Każdy z obszarów jest definiowany przez jego granice w przestrzeni cech i w przypadku, gdy zmienne x_1, x_2, \dots, x_M mają charakter ilościowy (są wskaźnikami finansowymi), można go przedstawić jako iloczyn:

$$I(\mathbf{x} \in R_k) = \prod_{m=1}^M I(v_{km}^{(d)} \leq x_m \leq v_{km}^{(g)}) \quad (25)$$

gdzie wartości $v_{km}^{(d)}$ oraz $v_{km}^{(g)}$ oznaczają odpowiednio górną i dolną granicę odcięcia w m -tym wymiarze przestrzeni, a I jest funkcją wskaźnikową:

$$I(q) = \begin{cases} 1 & \text{gdy } q \text{ jest prawdziwe,} \\ 0 & \text{w przeciwnym przypadku} \end{cases} \quad (26)$$

Problem oddzielenia (odseparowania) obiektów należących do różnych klas jest rozważany na podstawie statystycznej analizy wielowymiarowej w ramach zagadnienia klasyfikacji. W tym celu konieczne jest posiadanie zbioru obiektów S , nazywanego zbiorem uczącym, w którym przynależność do określonej klasy jest znana, tj. wartość zmiennej y dla każdego z nich została poprawnie zaklasy-

* (Gatnar, Walesiak, 2004, s. 107).

fikowana. Celem analizy jest znalezienie charakterystyk klas, tak aby wykorzystać je do klasyfikacji zbioru rozpoznawanego.

Zmienna y w modelu (24) jest cechą nominalną, wobec tego model ten jest modelem klasyfikacyjnym, a parametry a_k tego modelu są wyznaczone zgodnie z zasadą majoryzacji:

$$a_k = \arg \max_l \{p(l|k)\} \quad (27)$$

gdzie $p(l|k)$ oznacza prawdopodobieństwo, że pewien obiekt z segmentu R_k należy do klasy l . Innymi słowy, formuła (27) mówi o tym, że w segmencie R_k zmienna y przyjmuje tę wartość l , która występuje najczęściej.

Model (24) jest konstruowany drogą krokowej procedury doboru zmiennych, a więc dzielenia przestrzeni \mathbf{X}^m hiperpłaszczyznami równoległymi do osi poszczególnych zmiennych. Klasyczna procedura rekurencyjnego podziału składa się z następujących kroków (Gatnar, Walesiak, 2004, s. 108):

1. Mając przestrzeń \mathbf{X}^m (w której znajduje się zbiór uczący S), sprawdź, czy jest ona jednorodna ze względu na wartości zmiennej zależnej y (lub spełniony został inny warunek stopu). Jeżeli tak, to zakończ pracę.
2. W przeciwnym razie rozważ wszystkie możliwe sposoby podziału przestrzeni \mathbf{X}^m na rozłączne segmenty R_1, \dots, R_K (z uwzględnieniem wartości kolejno wybieranych zmiennych objaśniających).
3. Dokonaj oceny jakości każdego z tych podziałów zgodnie z przyjętym kryterium homogeniczności i wybierz najlepszy z nich.
4. Podziel przestrzeń \mathbf{X}^m w wybrany sposób.
5. Wykonaj kroki 1-4 rekurencyjnie dla każdego z segmentów R_1, \dots, R_K .
6. Procedura podziału się kończy, jeżeli zostało osiągnięte jedno z kryteriów stopu. Najczęściej jest nim jednorodność obiektów w segmencie R_1, \dots, R_K lub (ze względów praktycznych) określona minimalna liczba obiektów w uzyskanych segmentach (podzbiorach).

Przebieg procesu podziału rekurencyjnego najlepiej reprezentuje graf spójny bez cykli, czyli drzewo. Stąd wynika popularna nazwa tej metody – drzewa decyzyjne, a w sytuacji gdy zmienna y ma charakter jakościowy – drzewa klasyfikacyjne. Metody podziału rekurencyjnego oraz ich prezentacja w postaci drzew decyzyjnych są stosowane do wykrywania związków i relacji występujących w dużych zbiorach danych. Są więc one technikami stosunkowo młodej dyscypliny naukowej, jaką jest *Data Mining**

* W literaturze polskiej wyczerpujące informacje dotyczące różnych wariantów metody drzew klasyfikacyjnych i regresyjnych znaleźć można m.in. w pracach: (Gatnar, 2001; Gatnar, 2008 lub Gatnar, Walesiak, 2004).

5. Wybrane przykłady zastosowań przedstawionych metod do przewidywania bankructwa

W niniejszym punkcie zostaną przytoczone pierwsze zastosowania na świecie prezentowanych powyżej metod jako narzędzi predykcji upadłości firm oraz reprezentatywne przykłady ich zastosowania w prognozowaniu bankructwa firm polskich.

Pierwszym, który w 1968 r. wykorzystał liniową funkcję dyskryminacyjną Fishera dla celów prognozowania bankructwa był Altman (1968). Wyodrębnił on dwie grupy firm amerykańskich: bankrutów i niebankrutów. Jego próba wynosiła 66 korporacji, spośród których 33 stanowiły przedsiębiorstwa, jakie zbankrutowały w latach 1946-1965, a 33 to wylosowane firmy działające na rynku amerykańskim, które nie zbankrutowały w badanym okresie. Źródłem danych były sprawozdania finansowe tych przedsiębiorstw (dla bankrutów z roku poprzedzającego bankructwo). Wstępnie rozpatrywał on 22 wskaźniki finansowe reprezentujące klasyczne ich grupy: płynność finansową, rentowność, wspomaganie finansowe, wypłacalność i aktywność ekonomiczną. Po wstępnej analizie ekonomicznej, opartej na doświadczeniu profesjonalnym audytorów, do modelu wybranych zostało 5 wskaźników finansowych.

Klasyczny model Altmana, nazwany przez autora „Z-Score model” jest następujący (Za: Pocięcha, 2006, s. 206):

$$Z = 1,2 X_1 + 1,4 X_2 + 3,3 X_3 + 0,6 X_4 + 1,0 X_5$$

gdzie:

X_1 – kapitał pracujący / majątek ogółem,

X_2 – zysk zatrzymany / majątek ogółem,

X_3 – zysk przed opodatkowaniem (EBIT) /majątek ogółem,

X_4 – wartość rynkowa kapitału akcyjnego / wartość księgowa zadłużenia,

X_5 – przychody ze sprzedaży / majątek ogółem.

Wartość krytyczna (progowa) wyniosła 2,675. Według badań Altmana zdolność poprawnej klasyfikacji w prezentowanym modelu wynosiła 95%, przy czym błąd klasyfikacji I typu (zaklasyfikowanie bankruta jako niebankruta) wynosił 6%, a błąd II typu (zaklasyfikowanie niebankruta jako bankruta) – 3%. W rezultacie badań dotyczących minimalizacji prawdopodobieństwa błędnej klasyfikacji, Altman przyjął następujące wartości graniczne (Z) dla prognozowania bankructwa:

1,81 lub mniej – duże prawdopodobieństwo bankructwa firmy (strefa I – wszystkie firmy w próbie zbankrutowały),

3,00 lub więcej – niskie prawdopodobieństwo bankructwa firmy (strefa II – żadna firma nie zbankrutowała),

$1,81 < Z < 2,67$ – niepewna przyszłość firmy.

Model ten, mający już historyczne znaczenie, stał się punktem odniesienia do wszystkich późniejszych modeli prognozowania bankructwa. Altman później wielokrotnie modyfikował prezentowane przez siebie modele i testował je na różnych próbach firm*.

Pierwszym przykładem zastosowania metody dyskryminacyjnej do prognozowania bankructwa polskich firm jest praca Mączyńskiej (1994), natomiast za najbardziej rozbudowany przykład zastosowania wielowymiarowej analizy dyskryminacyjnej do prognozowania upadłości w polskich realiach gospodarczych można uznać pracę Hołdy (2006). Zostanie przedstawiony przykład liniowej funkcji dyskryminacyjnej zbudowanej i szacowanej przez autora niniejszego artykułu. Model został zbudowany na podstawie wskaźników płynności i rentowności.

W literaturze przyjmuje się, że zdolność klasyfikacyjna i predyktywna modelu dyskryminacyjnego jest tym lepsza, im więcej obszarów analizy finansowej reprezentują zmienne włączane do modelu. Punktem wyjścia do konstrukcji tego modelu był zbiór 13 zmiennych dyskryminujących. Ich dobór do modeli dyskryminacyjnych odbywał się drogą selekcji krokowej z poszczególnych grup wskaźników. W rezultacie otrzymano następujący model zbudowany na podstawie wskaźników płynności i rentowności:

$$Y_1 = -1,843 - 0,486 \text{ ROE} + 1,116 \text{ WKOP} - 0,500 \text{ WPB}$$

gdzie:

ROE – rentowność kapitału własnego,

WKOP – wskaźnik kosztów operacyjnych,

WPB – wskaźnik płynności bieżącej.

Model ten charakteryzuje się 92,50-proc. zdolnością do poprawnej klasyfikacji oraz 86,25-proc. zdolnością predyktywną. W próbie podstawowej błąd I rodzaju wynosi 6,67% (4/56), a błąd II rodzaju – 8,33% (5/55). W próbie testowej błąd I rodzaju wynosi 22,50% (9/31), a błąd II rodzaju – 5,00% (2/38). Wystarczyło przyjęcie trzech wymienionych zmiennych dyskryminujących, gdyż włączenie któregośkolwiek z pozostałych wskaźników płynności czy rentowności nie poprawiało sprawności modelu.

Pierwszym na świecie, który zastosował model logitowy do prognozowania bankructwa był Ohlson (1980). Zmienne do swojego modelu wybierał, opierając się na wskazaniach literatury z zakresu analizy finansowej. W efekcie uwzględ-

* Podsumowanie doświadczeń z budową i zastosowaniami zaproponowanych modeli zawarł on w pracy: Altman: Predicting Financial Distress of Companies: Revisiting the Z-Score and ZETA® Models, <http://pages.stern.nyu.edu/~ealtman/Zscores.PDF>.

nił w modelu 9 wskaźników finansowych. Jego próba składała się ze 105 bankrutów i 2058 dobrze funkcjonujących firm w latach 1970-1976. Model Ohlsona (w formie zlinearyzowanej) przedstawiał się następująco:

$$Y = -1,3 - 0,4 X_1 + 6,0 X_2 - 1,4 X_3 + 0,1 X_4 - 2,4 X_5 - 1,8 X_6 + 0,3 X_7 - 1,7 X_8 - 0,5 X_9$$

gdzie:

X_1 – log (aktywa ogółem /ogólny indeks cen),

X_2 – zobowiązania ogółem/aktywa ogółem,

X_3 – kapitał pracujący/aktywa ogółem,

X_4 – zobowiązania bieżące/aktywa obrotowe,

$X_5 = 1$, gdy zobowiązania ogółem przekraczają aktywa ogółem, 0 w przeciwnym przypadku;

X_6 – zysk netto/aktywa ogółem,

X_7 – przychody/zobowiązania ogółem,

$X_8 = 1$, gdy firma przynosiła straty w ciągu ostatnich dwóch lat, w przeciwnym przypadku 0,

X_9 – miara zmian zysku netto.

Od tego momentu nastąpił gwałtowny rozwój zastosowań modelu logitowego dla celów prognozowania bankructwa. W Polsce jako pierwsze zastosowanie modelu logistycznego do predykcji upadłości polskich firm przedstawione zostało w pracy Hołdy (2000).

Spośród wielu publikacji z tego obszaru w Polsce należy zwrócić uwagę na wyniki uzyskane przez Wędzkiego (2005). Przedstawił on wiele modeli logitowych upadłości w gospodarce polskiej. Przykładowo zaprezentowano tutaj model M3 – wielogałęziowy model dla przedsiębiorstw przemysłowych:

$$Y = 1,0 - 5,0 WB + 4,721 UKON + 3,598 WZO - 0,334 WUO + 0,048 IDF + 0,021 CN + 0,061 RIR$$

gdzie:

WB – bieżący wskaźnik płynności,

UKON – udział kapitału obrotowego w aktywach,

WZO – wskaźnik zadłużenia ogólnego,

WUO – wskaźnik udziału odsetek,

IDF – indeks dźwigni finansowej,

CN – cykl należności,

RIR – stopa zysku rezydualnego.

Próba przedsiębiorstw wynosiła: 40 bankrutów i 40 firm dobrze prosperujących.

Obszerny przegląd modeli regresji logistycznej oszacowanych dla celów prognozowania upadłości firm polskich jest zawarty m.in. w pracy Prusaka (2005).

Sieci neuronowe jako narzędzie prognozowania bankructwa były wykorzystywane w świecie od lat 90. ubiegłego wieku*. Szczególnie interesująca jest praca Odoma i Shardy (1990). Autorzy podjęli próbę zastosowania sieci neuronowych dla 5 wskaźników finansowych uwzględnionych w klasycznym, dyskryminacyjnym modelu Altmana prognozowania bankructwa. Posłużyli się oni danymi finansowymi, dotyczącymi 128 firm, pochodzących z okresu na rok przed bankructwem. Zdolność do poprawnej klasyfikacji przez model neuronowy dla firm „bankrutów” mieściła się w granicach 77,8%-81,5%, a dla firm „niebankrutów” – 78,6%-58,7%.

Jako pierwszy w Polsce propozycję zastosowania sztucznych sieci neuronowych dla celów prognozowania bankructwa polskich firm przedstawił Michaluk (2000). Przyjął on zasadę, że liczba neuronów w warstwie ukrytej nie może być większa od liczby neuronów w warstwie wejściowej. Do swojego badania wykorzystał dane dotyczące 259 przedsiębiorstw, w tym 79 bankrutów oraz 180 firm o dobrej kondycji finansowej. Rozpatrywał on 12 podstawowych wskaźników finansowych, charakteryzujących kondycję finansową przedsiębiorstw. Podstawowe parametry konstrukcji tej sieci były następujące:

- 12 neuronów w warstwie wejściowej (wskaźniki finansowe X_1, \dots, X_{12}),
- 12 neuronów w warstwie ukrytej,
- 1 neuron wyjściowy (RB – bankrut = 0; NBR – niebankrut = 1),
- sieć była uczona algorytmem wstecznej propagacji błędu.

Autor sieci przyjął trzy próby, różniące się ze względu na proporcje: niebankrut/bankrut, przyjmując odpowiednio: proporcje 1:1 (próba ucząca 40/40, próba testowa 39/39), 3:1 oraz 10:1 (próba ucząca 100/10; próba testowa 69/69). Zdolność do poprawnej klasyfikacji nie była całkiem zadowalająca, gdyż najmniejszy błąd I typu (procent bankrutów, które zostały nieprawidłowo zakwalifikowane do zbioru firm kontynuujących działalność) wynosił 10%, natomiast błąd drugiego typu (procent firm które kontynuują działalność, nieprawidłowo zakwalifikowanych do zbioru bankrutów) – 2%. Ogólna zdolność do poprawnej klasyfikacji (procent prawidłowo zaklasyfikowanych firm) była wysoka i wynosiła 97,27%, jednak generowała stosunkowo wysokie błędy I typu. W praktyce przyjmuje się, że koszty, jakie są ponoszone przez błędy I typu (uznanie firm, które później zbankrutowały za firmy zdrowe) są o wiele wyższe niż koszty błędów II typu.

Bogate wyniki badań prognostycznych wykonanych przy użyciu sieci neuronowych zaprezentowane zostały w pracy Korola i Prusaka (2005). Do kon-

* Pierwsze prace z tego zakresu to: (Bell, Ribar, Verchio, 1990 oraz Odom, Sharda, 1990).

strukcji modelu sieci neuronowej wykorzystano dane finansowe dotyczące 180 polskich przedsiębiorstw produkcyjnych z lat 1998-2001. Próba ucząca składała się z 39 firm będących w dobrej kondycji oraz 39 firm zagrożonych upadłością. Firmy znajdujące się w dobrej sytuacji wybierano drogą parowania z potencjalnymi bankrutami, tak aby one pochodziły zawsze z tej samej branży. Próby uczące konstruowano dla danych na rok i na dwa lata przed złożeniem wniosku o upadłość. Próba testowa również składała się z 39 przedsiębiorstw zdrowych oraz 39 bankrutów. Do budowy sieci neuronowej autorzy przyjęli 27 wskaźników finansowych charakteryzujących płynność finansową, zadłużenie, sprawność działania i rentowność firm. Dodatkowo autorzy wprowadzili zmienną pozafinansową, jaką był region działania danej spółki. Wykorzystano tutaj mapę intensywności zagrożenia upadłością firm w poszczególnych województwach. Sieć miała więc pierwotnie 28 sygnałów wejściowych, których zbiór autorzy oznaczyli jako K1. Alternatywnie zaproponowano redukcję sygnałów wejściowych (wskaźników finansowych) na podstawie wyników analizy współczynników korelacji pomiędzy przyjętymi pierwotnie cechami. Doprowadziło to do redukcji sygnałów wejściowych do czterech. Zbiór ten oznaczono jako K2. W trzecim wariancie (K3) przyjęto pięć wskaźników z modelu Altmana. Przyjmowano także wiele wariantów architektury sieci, różniących się liczbą neuronów w warstwie ukrytej oraz dla prób uczących i testowych na rok oraz na dwa lata przed upadłością. Rozważano również różne proporcje niebankrut/bankrut w strukturze prób.

Wyniki badań potwierdziły, że „zachodnie” wskaźniki finansowe, jakie były przyjmowane m.in. w modelu Altmana, nie sprawdzają się w warunkach gospodarki polskiej, gdyż podejście K3 dawało najgorsze rezultaty. Natomiast efektywność prognostyczną pozostałych modeli badano na próbach testowych. Skuteczność prognostyczna (procent prawidłowo zaklasyfikowanych firm w zbiorze testowym) była następująca: w wariancie K1 na rok przed upadkiem firmy wynosiła 98,72%, a na dwa lata przed upadkiem – 91,03%. W wariancie K2 na rok przed upadkiem wynosiła 97,44% a na dwa lata przed upadkiem – 87,18%. Jak wynika z powyższego, najlepsze rezultaty otrzymano dla najbardziej rozbudowanej sieci neuronowej.

Metodologia drzew klasyfikacyjnych bardzo wcześnie została zastosowana dla celów prognozowania bankructwa. Pierwszą pracą z tego zakresu był artykuł Frydman, Altmana i Kao (1985). Autorzy zastosowali algorytm RPA (*Recursive Partitioning Algorithm*) zaproponowany przez Breimana i in. (1984) do badania zdolności kredytowej i upadłości firm. Jedno z przedstawionych tam drzew klasyfikacyjnych stworzonych przez Frydman, Altmana i Kao charakteryzowało się wysoką zdolnością do poprawnej klasyfikacji ogółem, wynoszącą 94,00%. Praca

Frydman Altmana i Kao stała się później punktem odniesienia do badania efektywności innych metod prognozowania bankructwa (Por. Barniv, McDonald, 1999).

Pierwszym, który w badaniach upadłości polskich firm zastosował metodologię drzew decyzyjnych w wersji algorytmu CART, był Hołda (2006). Zbudował on drzewo decyzyjne dla firm branży produkcyjnej oraz porównał zdolność do poprawnej klasyfikacji oraz zdolność prognostyczną zbudowanego drzewa klasyfikacyjnego z siecią neuronową zbudowaną dla tego samego zestawu wskaźników finansowych. Zdolność do poprawnej klasyfikacji (na zbiorze uczącym) wynosiła 93,24%, a na zbiorze testowym – 83,87%. Zbudowana dla tych samych wskaźników finansowych sieć neuronowa miała następującą efektywność: zdolność do poprawnej klasyfikacji na zbiorze uczącym wynosiła 92,57% a na zbiorze testowym – 77,42%. Drzewo klasyfikacyjne okazało się więc nieco bardziej efektywnym narzędziem przewidywania bankructwa niż sieć neuronowa, dla tego samego zestawu wskaźników finansowych.

Drzewa klasyfikacyjne upadłości firm przynależnych do innych branż, budowane metodą CART, zostały przedstawione w pracy Hołdy i Pocięchy (2009). Dla firm branży budowlanej istotnymi okazały się dwa wskaźniki:

- rentowności – uwzględniający zysk/stratę netto zwiększoną o niepodzielny wynik z lat poprzednich oraz podatek dochodowy w stosunku do majątku ogółem,
- przepływów pieniężnych z działalności inwestycyjnej (środki pieniężne netto z działalności inwestycyjnej/ majątek ogółem).

Zdolność do poprawnej klasyfikacji tego drzewa była bardzo wysoka i wynosiła 93,75%, ale zdolność prognostyczna na grupie testowej była dość niska i wynosiła 65,00%.

Dla branży handlowo-usługowej drzewo klasyfikacyjne obejmowało następujące wskaźniki finansowe:

- zadłużenia (zysk/strata netto w stosunku do zobowiązań krótkookresowych),
- rentowności – uwzględniający zysk/stratę netto zwiększoną o niepodzielny wynik z lat poprzednich oraz podatek dochodowy w stosunku do majątku ogółem,
- rentowności – uwzględniający zysk/stratę netto w stosunku do kosztów ogółu działalności,
- zadłużenia – jako stosunek zobowiązań oraz rezerw do sumy bilansowej.

Zdolność do poprawnej klasyfikacji firm do grupy bankrutów i niebankrutów była wysoka i wynosiła 91,76%. Wyraźnie niższa była jednak zdolność prognostyczna drzewa na zbiorze testowym – wynosiła 67,86%.

Jak widać z przytoczonych przykładów, metody statystyki wielowymiarowej, a także komputerowe metody iteracyjne są często wykorzystywane w praktyce badań ekonomicznych nad problemami upadłości firm w Polsce i na świecie.

6. Uwarunkowania efektywności metod klasyfikacyjnych w prognozowaniu bankructwa

Z przedstawionego w poprzednim punkcie przeglądu wynika, że różni autorzy preferowali odmienne podejścia metodologiczne do prognozowania bankructwa. W związku z tym rodzi się pytanie, czy istnieją metody, które z reguły dają bardziej precyzyjne prognozy bankructwa niż inne. Z szerokich badań porównawczych przeprowadzonych na świecie (Patrz np. Bellovary, Giacomino, Akers, 2007), a także przeglądów prowadzonych przez autora (Pociecha, 2010a lub Pocięcha, Pawełek, 2011) wynika, że nie można mówić o metodach „lepszych” lub „gorszych” dla celów prognozowania upadłości firm.

Jako kryterium sprawności prognostycznej metod wymienionych w poprzednim punkcie przyjmuje się następujące miary zdolności do poprawnej klasyfikacji:

1. Sprawność I rodzaju – udział (procent) firm, które zbankrutowały, prawidłowo zakwalifikowanych przez model do zbioru bankrutów.
2. Błąd I rodzaju – procent bankrutów, które zostały nieprawidłowo zakwalifikowane do zbioru firm kontynuujących działalność.
3. Sprawność II rodzaju – procent firm kontynuujących swoją działalność (niebankrutów), prawidłowo rozpoznanych przez model.
4. Błąd II rodzaju – procent firm które kontynuują działalność, nieprawidłowo zakwalifikowanych do zbioru bankrutów.
5. Sprawność ogólna – procent prawidłowo zaklasyfikowanych firm,
6. Błąd ogólny – procent nieprawidłowo zaklasyfikowanych firm.

Najpopularniejsze polskie modele przewidywania upadłości firm (Pociecha, 2010, s. 56-57) miały na zbiorze uczącym sprawność ogólną dla modeli dyskryminacyjnych w granicach 78,6%-93,2%, dla modeli logitowych: 89,0%-91,9%, dla sieci neuronowych: 93,9%-96,2%, dla drzew klasyfikacyjnych: 91,8%-93,8%. W tym przeglądzie widać, że nieco bardziej precyzyjne wyniki otrzymano z sieci neuronowych, lecz w literaturze podawano także przykłady, gdzie linio-wa funkcja dyskryminacyjna może dawać lepsze rezultaty niż zbyt skomplikowana sieć neuronowa. Zdolność prognostyczną bada się na zbiorze testowym, który na ogół jest wyodrębnioną częścią pierwotnego zbioru danych. Tutaj też nie stwierdzono wyraźnych różnic w zdolności prognostycznej omawianych typów modeli. Reasumując, można stwierdzić, że precyzja prognozy bankructwa nie zależy od typu modelu prognostycznego.

Wobec braku rozstrzygnięć, jakiego typu modele są najodpowiedniejsze dla celów prognozowania bankructwa, należy sobie postawić fundamentalne pyta-

nie: Jakie są źródła błędów popełnianych w procesie prognozowania bankructwa? (Pociecha, 2011, s. 128-129).

Jednym z nich jest wartościowy charakter wskaźników finansowych. Istnieją krajowe i międzynarodowe standardy sprawozdawczości finansowej, ale daleko jeszcze do ujednoczenia sposobu pomiaru wielkości finansowych, szczególnie w skali międzynarodowej. Precyzja pomiaru wskaźników finansowych jako zmiennych klasyfikujących do zbioru bankrutów lub niebankrutów nie jest więc zbyt wysoka.

Drugie możliwe źródło błędów to metoda doboru prób. W klasycznym ujęciu np. metody dyskryminacyjnej, próby z badanych populacji są wybierane drogą losową. W praktyce doboru prób nie przeprowadza się w sposób nielosowy. Uwzględnia się na ogół wszystkie firmy upadłe w badanym okresie, a do niej nielosową metodą parowania dobiera się przedsiębiorstwa dobrze funkcjonujące. Nie można więc mówić o doborze losowym, w sensie klasycznym, a więc także o błędzie próbkowania. Testowany błąd klasyfikacji nie wynika z tego, że operujemy próbami losowymi.

Istotnym elementem błędów w prognozowaniu bankructwa jest tzw. bankructwo z przyczyn strategicznych. Zarządcy lub właściciele firmy dobrze prosperującej mogą celowo doprowadzić firmę do bankructwa, wyprowadzając nieco wcześniej jej aktywa np. do „rajów podatkowych”. Żaden model predykcji bankructwa nie uwzględnia celowego działania zarządców firm w kierunku celowego doprowadzenia do bankructwa.

Kolejnym źródłem błędów jest niestabilny charakter badanych populacji. Populacje bankrutów i przedsiębiorstw dobrze funkcjonujących w sytuacji koniunktury gospodarczej nie są identyczne z tymi populacjami w okresie kryzysu gospodarczego. Błąd prognozy może więc zależeć od tego, że model zbudowany został dla danych z okresu koniunktury, a prognoza budowana jest dla firmy w okresie recesji.

Podsumowując, warto jeszcze raz zwrócić uwagę na pionierski charakter pracy prof. J. Kolonki (1980), która w środowisku polskich statystyków przez lata motywowała do rozwijania teorii i zastosowań metod klasyfikacji danych, czego wyrazem jest ich użyteczność m.in. dla celów prognozowania upadłości firm.

Bibliografia

- Altman E.I. (1968): *Financial Ratios, Discriminant Analysis and Prediction of Corporate Bankruptcy*. „The Journal of Finance”, Vol. 23, September.
- Altman E.I. (2000): *Predicting Financial Distress of Companies: Revisiting the Z-Score and ZETA® Models*, <http://pages.stern.nyu.edu/~ealtman/Zscores.PDF>.
- Aziz M.A., Dar H.A. (2004): *Predicting Corporate Bankruptcy, Whither do We Stand?* 3rd Annual Meeting of the European Economics and Finance Society “Word Economy and European Integration”, University of Gdańsk, 13-16 May.
- Barniv R, McDonald J.B. (1999): *Review of Categorical Models for Classification Issues In Accounting and Finance*. „Review of Quantitative Finance and Accounting”, 13.
- Bell T.B., Ribar G.S., Verchio J. (1990): *Neural Nets Versus logistic Regression: A Comparison of Each Model's Ability to Predict Commercial Bank Failures*. In: *Proceedings of the 1990 Deloitte and Touché/University of Kansas Symposium of Auditing Problems*. Ed. R.P. Srivastava.
- Bellovary J., Giacomino D., Akers M. (2007): *A Review of Bankruptcy Prediction Studies: 1930 to Present*. „Journal of Financial Education”, Vol. 33, Winter.
- Breiman L., Friedman J., Olshen R., Stone C. (1984): *Classification and regression trees*. CRC Press, London.
- Christensen R. (1991): *Linear Models for Multivariate, Time Series, and Spatial Data*. Springer, New York.
- Fisher R.A. (1936): *The Use of Multiple Measurements in Taxonomic Problems*. „Annals of Eugenics”, 7.
- Frydman H., Altman E.I., Kao D. (1985): *Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress*. „Journal of Finance”, Vol. 40.1.
- Gatnar E. (2001): *Nieparametryczna metoda dyskryminacji i regresji*. Wydawnictwo Naukowe PWN, Warszawa.
- Gatnar E. (2008): *Podejście wielomodelowe w zagadnieniach dyskryminacji i regresji*. Wydawnictwo Naukowe PWN, Warszawa.
- Gatnar E., Walesiak M. (2004): *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, Wyd. AE, Wrocław.
- Giri N.C. (1996): *Multivariate Statistical Analysis*. Dekker, New York.
- Hołda A. (2000): *Optymalizacja i model zastosowania procedur analitycznych w rewizji sprawozdań finansowych*. Praca doktorska, Akademia Ekonomiczna, Kraków.
- Hołda A. (2006): *Zasada kontynuacji działalności i prognozowanie upadłości w polskich realiach gospodarczych*. Wydawnictwo Akademii Ekonomicznej, seria specjalna nr 174, Kraków.

- Hołda A., Pociecha J. (2009): *Probabilistyczne metody badania sprawozdań finansowych*. Wydawnictwo Uniwersytetu Ekonomicznego, Kraków.
- Kolonko J. (1980): *Analiza dyskryminacyjna i jej zastosowania w ekonomii*. PWN, Warszawa.
- Korbicz J., Obuchowicz A., Uciński D. (1994): *Sztuczne sieci neuronowe. Podstawy i zastosowania*. Akademicka Oficyna Wydawnicza PLJ, Warszawa.
- Korol T., Prusak B. (2005): *Upadłość przedsiębiorstw a wykorzystanie sztucznej inteligencji*. CeDeWu, Warszawa.
- Lula P. (1999): *Jednokierunkowe sieci neuronowe w modelowaniu zjawisk ekonomicznych*. Wydawnictwo AE, Kraków.
- Mączyńska E. (1994): *Ocena kondycji przedsiębiorstwa (Uproszczone metody)*. „Życie gospodarcze”, nr 38.
- McCulloch C.E., Searle S.R., Neuhaus J.M. (2009): *Generalized, Linear, and Mixed Models*. Wiley, New York.
- Michaluk K. (2000): *Efektywność modeli prognozujących upadłość przedsiębiorstw*. Praca doktorska, Uniwersytet Szczeciński, Szczecin.
- Odom M.D., Sharda R. (1990): *A Neural Network Model for Bankruptcy Prediction*. Proceedings of IEEE International Conference on Neural Networks, Vol. 2, San Diego.
- Ohlson J. (1980): *Financial Ratios and the Probabilistic Prediction of Bankruptcy*. „Journal of Accounting Research”, Spring.
- Ossowski S. (1996): *Sieci neuronowe w ujęciu algorytmicznym*. WNT, Warszawa.
- Pociecha J. (2006): *Funkcja dyskryminacyjna jako narzędzie prognozowania bankructwa firmy – metoda oraz rezultaty praktyczne*. W: *Matematyka język uniwersalny*. Wyd. AE, Kraków.
- Pociecha J. (2007): *Problemy prognozowania bankructwa firmy metodą analizy dyskryminacyjnej*. „Acta Universitatis Lodzianis, Folia Oeconomica”, nr 205.
- Pociecha J. (2010a): *Metodologiczne problemy prognozowania bankructwa*. „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, „Taksonomia 17, Klasyfikacja i analiza danych – teoria i zastosowania”, Wrocław.
- Pociecha J. (2010b): *Zastosowania sztucznych sieci neuronowych w prognozowaniu bankructwa firm*. W: *Nauki ekonomiczne wobec wyzwań współczesnej gospodarki światowej*. Uniwersytet Ekonomiczny, Kraków.
- Pociecha J. (2011): *Modele prognozowania bankructwa w systemie wczesnego ostrzegania przedsiębiorstw*. W: *Spoleczna rola statystyki*. Red. W. Ostasiewicz. „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu”, nr 165, Wrocław.
- Pociecha J. (2012): *Model logitowy jako narzędzie prognozowania bankructwa. Jego zalety i ograniczenia*. W: *Spotkania z królową nauk*. Wydawnictwo Uniwersytetu Ekonomicznego, Kraków.

- Pociecha J., Pawełek B. (2011): *Prognozowanie bankructwa a koniunktura gospodarcza*. „Metody analizy danych”, Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie, nr 873, Kraków.
- Prusak B. (2005): *Nowoczesne metody prognozowania zagrożenia finansowego przedsiębiorstwa*. Difin, Warszawa.
- Rencher A.C. (1998): *Multivariate Statistical Inference and Applications*. Wiley, New York.
- Tadeusiewicz R. (1993): *Sieci neuronowe*. Wydawnictwo Naukowe PWN, Warszawa.
- Wędzki D. (2005): *Bankruptcy Logit Model for Polish Economy*. „Argumenta Oeconomica Cracoviensia”, nr 3.
- Witkowska D. (2002): *Sztuczne sieci neuronowe i metody statystyczne. Wybrane zagadnienia finansowe*. C.H. Beck, Warszawa.

SELECTED CLASSIFICATION METHODS AND THEIR EFFECTIVENESS IN FIRMS' COLLAPSING PREDICTION

Summary

Classification methods are recognised as useful tool for bankruptcy prediction. Among them the most popular are: linear discriminant function, Logit model, neural network and classification tree. The ideas and basic formulas of these methods are presented in the paper. Some examples of application those procedures, which were published in world and Polish literature, are mentioned in the following parts of the paper. Some effectiveness conditions of presented methods are discussed. In the conclusion it has been stressed, that the precision of bankruptcy prediction not strictly depend on classification method which has been used. Sources of errors in bankruptcy prediction has been discussed on the end of the paper. Among them important are: valuated character of financial ratios, as an impute variables in such models, problems in samples selection, which usually hasn't random character and unstable character of considered populations. Probability of firms' collapse strongly depends on the stage of business cycle.