

Henryk Marjak

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

NEURONOWE MODELE IDENTYFIKOWANIA SYTUACJI FINANSOWEJ PRZEDSIĘBIORSTW

NEURONAL MODELS TO IDENTIFICATION FINANCIAL DISTRESS OF ENTERPRISES

Słowa kluczowe: sieci neuronowe, bankructwo, klasyfikacja

Key words: neural networks, business failure, classification

Synopsis. Przedstawiono badania dotyczące konstruowania modelu oceny sytuacji finansowej przedsiębiorstw w oparciu o sztuczne sieci neuronowe. Parametry sieci neuronowej optymalizowane były za pomocą algorytmu gradientów sprzężonych z regularyzacją oraz RPROP (resilient backpropagation). W zaproponowanych modelach sytuacja finansowa badana była na podstawie informacji, zawartych w raportach kwartalnych spółek branży spożywczej notowanych na WGPW.

Wstęp

Mimo, że tradycyjne modele prognozowania bankructwa opracowane w latach 60-70. nie są całkowicie satysfakcjonujące, to jednak mają zastosowanie dla grup przedsiębiorstw odpowiadających przedsiębiorstwom, na podstawie których modele te zostały opracowane. Rozwój technologii informacyjno-komunikacyjnych (ICT) stworzył jednakże możliwości bardziej elastycznego budowania modeli prognozowania sytuacji finansowej przedsiębiorstw.

Spośród klasycznych modeli predykcji bankructwa, nazywanych również systemami wczesnego ostrzegania, łączącymi wskaźniki finansowe wymienić należy model Beavera [1968], Altmana [1968] oraz Altmana, Holdemana i Narayanana [1977]. Te pierwsze modele bankructwa posiadały różne poziomy sukcesu w klasyfikacji bankrut/niebankrut. Badacze stosowali takie techniki klasyfikacyjne jak: wielowymiarowa analiza dyskryminacyjna lub analiza *logit/probit* badając dwie grupy firm: zdrowe i zagrożone. Na tej podstawie tworzony był model klasyfikacyjny, stosowany następnie do prognozowania nowych przypadków. Wyniki uzyskane za pomocą takich modeli było korzystne wtedy, gdy warunki środowiska ekonomicznego dla próbek w trakcie prognozy, pokrywały się z warunkami ekonomicznymi panującymi, gdy model był tworzony.

Pierwsze badania dotyczące zastosowania sieci neuronowych w ocenie zdolności kredytowej, przeprowadził również Altman, Marco i Vedetto [1994]. West [2000] opracował algorytmy sieci neuronowych, za pomocą których uzyskał wyniki lepsze niż wyniki uzyskane za pomocą np. drzew decyzyjnych w ocenie zdolności kredytowej. Sieci neuronowe dają dobre wyniki w przedsiębiorstwach działających w trudnych warunkach ekonomicznych, tj. bez rozwiniętej infrastruktury, z wysokimi kosztami transakcyjnymi i ograniczeniami prawnymi.

Wyniki porównania tradycyjnych metod statystycznych i metod alternatywnych wykazują zróżnicowanie, które zależy od uprzedzeń badaczy, na co zwrócili uwagę Liou i Yang [2008]. Badania przeglądowe Balcaena i Oogghe [2004] pokazały, że w 15 przypadkach (na 28) sieci neuronowe były „zdecydowanie najlepsze”, „warunkowo najlepsze” lub „nieco lepsze”.

Celem pracy było skonstruowanie takiego modelu, który umożliwiłby właściwą ocenę przedsiębiorstwa, a w rezultacie podjęcie optymalnych decyzji inwestycyjnych, jak też decyzji dotyczących współpracy z ocenianym przedsiębiorstwem.

Material i metodyka badań

Perceptron wielowarstwowy to adaptacyjna i zasadniczo nieliniowa maszyna licząca, zbudowana z połączenia prostych elementów przetwarzających, nazywanych neuronami. Liczba elementów i wzorzec połączeń między nimi definiuje strukturę sieci neuronowych. Sygnały przepływające

przez połączenia skalowane są przez odpowiednie parametry nazywane wagami synaptycznymi. Każdy element przetwarzający ma kilka połączeń wejściowych i tylko jedno wyjściowe.

Proces kończący się zastosowaniem i skutecznym wykorzystaniem sieci neuronowej poprzedzony powinien być, oprócz procesu który ogólnie nazwać możemy uczeniem sieci neuronowej, także pracami przygotowawczymi. Realizację procesu badania sieci neuronowej podzielić można na następujące etapy:

- określenie badanej wielkości,
- gromadzenie danych, szeregów liczbowych opisujących badane zjawisko,
- analiza danych, wstępne ich przetworzenie (*preprocessing*),
- wyodrębnienie zbiorów uczącego i testowego,
- wybór architektury sieci i metody uczenia,
- uczenie sieci neuronowej,
- weryfikacja modelu,
- zastosowanie modelu.

W pracy zastosowano dwa algorytmy uczenia gradientów sprzężonych z regularyzacją (SCG) oraz Rprop, dla obu tych metod proces uczenia można uogólnić wyróżniając kilka etapów:

- analiza sieci neuronowej dla aktualnego wektora wejściowego,
- utworzenie sieci propagacji wstecznej, zastąpienie funkcji aktywacji ich pochodnymi, podanie na bieżące wyjścia obecne wejścia, analiza różnic między odpowiednimi wartościami uzyskanymi i wymaganymi,
- adaptacja wag na podstawie poprzednich etapów,
- kontynuacja uczenia dla wszystkich wektorów uczących do czasu spełnienia warunku zatrzymania.

Warunek zatrzymania to, albo uzyskanie założonego błędu, np. MSE (błąd średniokwadratowy), albo wykonanie zaplanowanej liczby cykli uczenia.

$$MSE = \frac{1}{n-p} \sum_{i=0}^{n-1} (t_i - \sigma_i)^2$$

gdzie:

t_i – wartości rzeczywiste,

σ_i – wartości otrzymane,

n – liczba próbek uczących,

p – liczba szacowanych parametrów (wagi).

Metoda gradientów sprzężonych z regularyzacją [Moller 1993] (SCG – *scaled conjugate gradient*) jest odmianą zwykłej metody gradientów sprzężonych. Ogólnie mówi się o tej grupie metod, jako o metodach drugiego stopnia. Określenie drugi stopień, oznacza w tym miejscu fakt używania drugiej pochodnej funkcji celu. Techniki te znajdują lepszą trasę do minimum, niż techniki wykorzystujące pierwszą pochodną, odbywa się to wyższym kosztem obliczeniowym.

Standardowa propagacja błędów zakłada obliczenia w kierunku zmniejszającego się gradientu funkcji błędów. Metody gradientów sprzężonych wykorzystują obliczenia sprzężone z poprzednim etapem obliczeniowym. W przypadku metody gradientów sprzężonych z regularyzacją, obowiązują następujące zależności:

$$s_k = E''(w_k) \cdot p_k \approx \frac{E'(w_k + \sigma_k p_k) - E'(w_k)}{\sigma_k}, 0 < \sigma_k \ll 1$$

$$s_k = \frac{E'(w_k + \sigma_k p_k) - E'(w_k)}{\sigma_k} + \lambda_k p_k$$

gdzie: $w_{i+1} = w_i + \alpha_i p_i$

w_k – wektor wag,

$E(w_k)$ – funkcja błędów,

k – etap iteracji,

p_j – nowy kierunek sprzężenia,

σ_k – wartość kroku w kierunku sprzężenia,

λ_k – parametr regularyzacji.

Metoda jest znacznie szybsza, w sensie dochodzenia do określonej wartości błędu, od podstawowych metod gradientów sprzężonych, jednakże zaleta ta związana jest z dłuższym czasem obliczeń. Mówiąc zatem o prędkości uczenia, nie mamy na myśli czasu wykorzystania procesora, ale ilość epok uczenia koniecznych do osiągnięcia minimum błędu.

Nazwa algorytmu RPROP [Riedmiller 1993] pochodzi od ang. *Resilient Backpropagation*. Podstawowym założeniem tej metody jest eliminacja szkodliwego wpływu wielkości pochodnych cząstkowych. Dlatego rozważany jest tylko znak pochodnej, na podstawie którego określany jest kierunek zmian wag.

Wielkość zmian wag wyznaczany jest za pomocą czynnika $\Delta_{ij}^{(t)}$:

$$\Delta w_{ij}^{(t)} = \begin{cases} -\Delta_{ij}^{(t)} & \text{gdy } \frac{\partial E^{(t)}}{\partial w_{ij}} > 0 \\ +\Delta_{ij}^{(t)} & \text{gdy } \frac{\partial E^{(t)}}{\partial w_{ij}} < 0 \\ 0 & \text{w innym przypadku} \end{cases}$$

$$\Delta_{ij}^{(t)} = \begin{cases} \eta^+ \cdot \Delta_{ij}^{(t-1)} & \text{gdy } \frac{\partial E^{(t-1)}}{\partial w_{ij}} \cdot \frac{\partial E^{(t)}}{\partial w_{ij}} > 0 \\ \eta^- \cdot \Delta_{ij}^{(t-1)} & \text{gdy } \frac{\partial E^{(t-1)}}{\partial w_{ij}} \cdot \frac{\partial E^{(t)}}{\partial w_{ij}} < 0 \\ \Delta_{ij}^{(t-1)} & \text{w innym przypadku} \end{cases}$$

gdzie:

$\frac{\partial E^{(t)}}{\partial w_{ij}}$ – gradient po wszystkich wzorcach uczących,

η^+, η^- – maksymalna i minimalna wartość współczynnika uczenia.

Algorytm ten znacznie przyspiesza proces uczenia w obszarach, gdzie nachylenie funkcji błędu jest niewielkie. Podstawą tej metody jest ciągły wzrost współczynnika uczenia, gdy dla dwu kolejnych kroków, znaki gradientu są takie same, a także zmniejszenie wartości współczynnika, gdy znaki te są różne.

Budowę modelu oparto na istniejących modelach prognozowania bankructwa, tj.:

1. Model wczesnego ostrzegania Beaver'a [1966]:

$$P_{ij} = f(\text{CASH}_i, \text{NIN}_i, \text{LEV}_i, \text{WC}_i, \text{LIQ}_i, \text{NCRE}_i)$$

2. Model wczesnego ostrzegania Altmana, [1968]:

$$P_{ij} = f(\text{WC}_i, \text{CP}_i, \text{ROA}_i, \text{CAP}_i, \text{ACT}_i)$$

3. Zmodyfikowany model wczesnego ostrzegania Altmana, Hodelmana i Narayanana [1977]:

$$P_{ij} = f(\text{ROA}_i, S_{\text{ROA}}, \text{DS}_i, \text{CP}_i, \text{LIQ}_i, \text{CAP}_i, \text{SIZE}_i)$$

4. Model Garcia, Arquesa i Calvo-Floresa [2007]:

$$P_{ij} = f(\text{QLIQ}_i, \text{ATD}_i, \text{INT}_i, \text{AMO}_i, \text{EAR}_i)$$

gdzie:

CASH_i – nadwyżka finansowa (wynik finansowy netto + amortyzacja)/zobowiązania ogółem, NIN_i – wynik finansowy netto/aktywa ogółem, LEV_i – zobowiązania ogółem/aktywa ogółem (wskaźnik zadłużenia ogólnego); WC_i – kapitał obrotowy/aktywa ogółem, LIQ_i – aktywa bieżące/zobowiązania krótkoterminowe, NCRE_i – (aktywa płynne – zobowiązania krótkoterminowe)/(wydatki na działalność operacyjną – amortyzacja), CP_i – zysk zatrzymany/aktywa ogółem, ROA_i – zysk przed opodatkowaniem i spłatą odsetek/aktywa ogółem, CAP_i – rynkowa wartość kapitału akcyjnego/księgowo wartość zadłużenia, ACT_i – przychody ze sprzedaży/aktywa ogółem, S_{ROA} – błąd standardowy oszacowania ROA, DS_i – zyski bez uwzględnienia odsetek i przed opodatkowaniem/odsetki ogółem, SIZE_i – transformacja logarytmiczna aktywów ogółem, QLIQ_i – aktywa płynne/płatności bieżące, ATD_i – aktywa ogółem/zobowiązania ogółem, INT_i – odsetki/sprzedaż, AMO_i – amortyzacja/aktywa amortyzowane, EAR_i – zyski przed opodatkowaniem/zobowiązania ogółem.

Ocenę klasyfikatorów przeprowadzono porównując macierze pomyłek (*confusion matrix*), w szczególności współczynnik pomyłek dla zbioru testowego, charakterystyki operacyjne odbiornika ROC (*Receiver Operating Characteristics*) oraz minimalizowane w trakcie uczenia błędy (MSE z regularyzacją i ekonomizacją). Klasyfikatory, w których zastosowano obie metody uczenia (SCG i Rprop), porównano za pomocą k -krotnego testu krzyżowego (*k-fold cross validation*).

Tabela 1. Macierz pomyłek

		Przewidywane klasy	
		pozytywne	negatywne
Oryginalne klasy	pozytywne	TP	FN
	negatywne	FP	TN

Źródło: opracowanie własne.

Macierz pomyłek to tabela zawierająca liczbę wierszy i kolumn odpowiadającą klasom decyzyjnym. Wiersze odpowiadają oryginalnym klasom decyzyjnym, kolumny przewidywanym klasom decyzyjnym (tab. 1). W macierzy pomyłek oznaczenia to: TP (*true positive*) – liczba poprawnie sklasyfikowanych pozytywnych przypadków, TN (*true negative*) – liczba poprawnie sklasyfikowanych negatywnych przypadków, FP (*false positive*) – liczba błędnie sklasyfikowanych pozytywnych przypadków, FN (*false negative*) – liczba błędnie sklasyfikowanych negatywnych przypadków. Na podstawie powyższych wartości określić można czułość (*sensitivity*), czyli $TP/(TP+FN)$ oraz specyficzność (*specifity*), tj. $TN/(TN+FP)$. Charakterystyka odbiornika ROC określa zależności pomiędzy czułością a specyficznością, określając jakość klasyfikatora. Im pole wyznaczone przez krzywe bliższe jest wartości 1, tym klasyfikator posiada lepsze właściwości klasyfikacyjne.

Porównanie klasyfikatorów za pomocą k -krotnego testu krzyżowego polega na podziale zbioru danych na k części. Model trenowany jest na $k-1$ częściach, następnie testowany na pozostałej części. Procedura powtarzana jest k -krotnie, w ten sposób każda z części jest użyta do testowania. W naszych rozważaniach przyjęto $k=10$. Wyniki uzyskane w ten sposób sprawdzono za pomocą odpowiednich testów statystycznych.

Dane, które poddano klasyfikacji pozyskano z raportów kwartalnych spółek giełdowych wchodzących w skład indeksu Wig-spożywczy¹ za lat 2001-2009 (łącznie 650 wektorów).

Analiza wyników

Zbiór danych, zgodnie z wcześniejszymi założeniami, podzielono na 10 części – 9 z nich służyło do trenowania, a jedna do testowania. Wektory wejściowe zawierały takie wskaźniki jak: rentowności – ROA, ROE, zadłużenia – LEV i aktywności – ACT. Jakość klasyfikacji oceniono za pomocą współczynnika pomyłek (*con_rp*, *con_scg*) oraz błędu MSE (*mse_rp*, *mse_scg*) obliczone dla każdej z części testowych. Tabela 2 zawiera statystyki opisowe wymienionych błędów.

Analiza błędów wskazuje, że dla algorytmu uczenia SCG uzyskano najmniejsze wartości tak współczynnika pomyłek, jak i MSE (odpowiednio: 0,3077 i 0,2418). Jednakże dla tej metody uczenia uzyskano większą, niż w przypadku algorytmu Rprop, wartość maksymalną współczynnika pomyłek (*con_rp*=0,4308 wobec *con_scg*=0,4769). Wartości odchylenia standardowego współczynni-

Tabela 2. Statystyki opisowe dla danych testowych

Wyszczególnienie	Mediana	Minimum	Maksimum	Dolny kwartyl	Górny kwartyl	Odchylenie standardowe
<i>con_rp</i>	0,3692	0,3231	0,4308	0,3538	0,4000	0,0341
<i>con_scg</i>	0,3615	0,3077	0,4769	0,3231	0,3846	0,0510
<i>mse_rp</i>	0,2692	0,2535	0,3455	0,2561	0,2802	0,0326
<i>mse_scg</i>	0,2675	0,2418	0,2998	0,2560	0,2848	0,0182

Źródło: opracowanie własne.

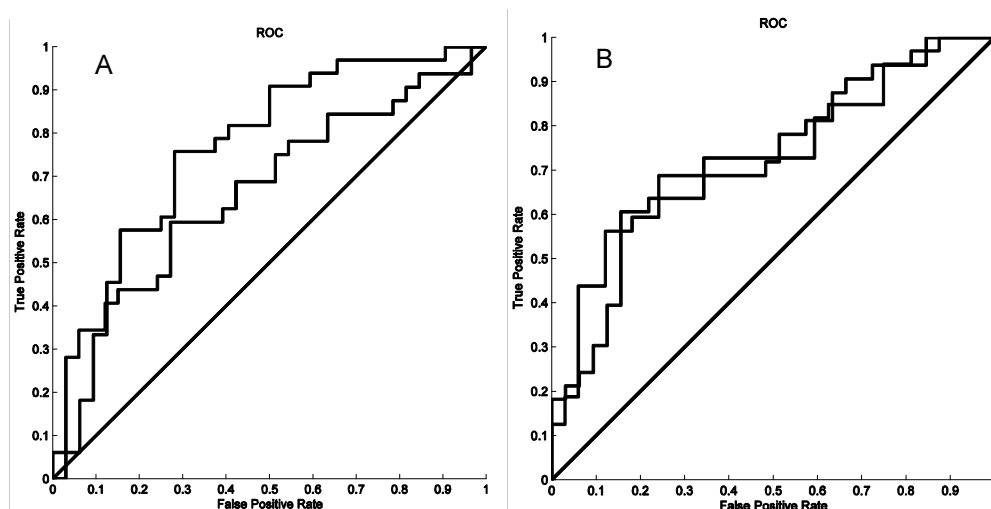
¹ Skład Wig-spożywczy (30.04.2010 r.): Kernel Holding SA, Kofola SA, Astarta Holding NV, Wawel SA, Jutrzenka SA, Elstar Oils SA, Polski Koncern Mięśny Duda SA, Indykpol SA, Mispol SA, Ambra SA, Graal SA, Pamapol SA, Zakłady Przemysłu Cukierniczego MIESZKO SA, Makarony Polskie SA, WILBO SA, Seko SA, Przedsiębiorstwo Przemysłu Spożywczego PEPEES SA, Belvedere SA.

Tabela 3. Macierze pomyłek dla Rprop i SCG

Oryginalne klasy	Przewidywane klasy		
	1	2	
1	18 27,7%	6 9,2%	75,0% 25,0%
2	15 23,1%	26 40,0%	63,4% 36,6%
	54,5% 45,5%	82,3% 18,8%	67,7% 32,3%

Oryginalne klasy	Przewidywane klasy		
	1	2	
1	21 32,3%	8 12,3%	72,4% 27,6%
2	12 18,5%	24 36,9%	66,7% 33,3%
	63,6% 36,4%	75,0% 25,0%	69,2% 30,8%

Źródło: opracowanie własne.



Rysunek 1. Charakterystyki ROC dla Rprop i SCG

Źródło: opracowanie własne.

ków pomyłek wskazują na mniejsze rozproszenie tych wartości w przypadku algorytmu Rprop.

Analiza miar położenia wskazuje na lepszą zdolność klasyfikacyjną modelu, który uczony był z wykorzystaniem algorytmu SCG, jednakże porównanie obu tych metod za pomocą k -krotnego sprawdzianu krzyżowego, nie wskazuje na istotne statystycznie zróżnicowanie wartości, tak współczynnika pomyłek (test t -Studenta), jak i MSE (test U Manna-Whitneya). Na rysunku 1 przedstawiono macierze pomyłek i charakterystyki ROC dla najlepszych modeli optymalizowanych za pomocą Rprop (tab. 3A, rys. 1A) i SCG (tab. 3B, rys. 1B).

Wnioski

Modele statystyczne identyfikacji sytuacji finansowej przedsiębiorstw wykazują dużą skuteczność, gdy prognozy dotyczą jednostek podobnych, do jednostek na podstawie których zostały zbudowane. W związku z tym, że budowane były na podstawie danych z przedsiębiorstw amerykańskich, trudno zastosować je w stosunku do przedsiębiorstw polskich. Modele te mogą charakteryzować się wyższą sprawnością w stosunku do przedsiębiorstw, działających w stabilnej do niedawna gospodarce amerykańskiej, niż w przedsiębiorstwach polskich działających w mniej przewidywalnych warunkach. Przeprowadzone badania wskazują, że modele nieliniowe takie, jak: sieci neuronowe, mogą być dobrą alternatywą wobec klasycznych modeli liniowych.

W pracy skupiono się na porównaniu dwóch algorytmów uczenia: gradientów sprzężonych z regularyzacją i Rprop. Dla zbadanych zbiorów nie stwierdzono istotnego statystycznie zróżnicowania jakości klasyfikatorów (10-krotny test krzyżowy) optymalizowanych tymi dwiema metodami.

W dalszych badaniach sugeruje się rozszerzenie zbioru danych o większą liczbę przypadków, a także i zmiennych. Metodologia może być rozszerzona o inne techniki sztucznej inteligencji takie, jak: algorytmy genetyczne, systemy eksperckie i maszyny wektorów nośnych.

Literatura

- Altman E.I., Haldeman E., Narayanan P.** 1977: Z analysis: A new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance*, vol. 10, s. 29-54.
- Altman E.I., Marco G., Varetto F.** 1994: Corporate Distress Diagnosis: Comparison Using Linear Discriminant Analysis and Neural Networks (the Italian Experience). *Journal of Banking and Finance*, nr 18, s. 505-529.
- Altman I.E.** 1968: Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, september, s. 589-609.
- Balcaen S., Ooghe H.** 2004: Alternative methodologies in studies on business failure: do they produce better results than the classic statistical methods? Vlerick Leuven Gent Management School, Gent. *Working Paper Series*, nr 16,
- Beaver W.H.** 1968: Alternative accounting measures as predictor of failure. *The Accounting Review*, Autumn, s. 112-22.
- Becerra V.M., Galvão R.K.H., Abou-Seada M.** 2005: Neural and Wavelet Network Models for Financial Distress Classification. *Data Mining and Knowledge Discovery*, nr 11, s. 35-55.
- Garcia D., Arques A., Calvo-Flores A.** 1997: Un modelo discriminante para evaluar el riesgo bancario en los créditos a empresas. *Revista Espanola de Financiacion y Contabilidad*, vol. 24 No. 82, s. 175-200.
- Liou F.M., Yang C.H.** 2008: Predicting business failure under the existence of fraudulent financial reporting. *International Journal of Accounting and Information Management*, vol. 16 No. 1, s. 74-86.
- Moller M.F.** 1993: A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, vol. 6, s. 525-533.
- Riedmiller M., Braun H.** 1992: A fast adaptive learning algorithm. Technical Report, University Karlsruhe, Germany.
- West D.** 2000: Neural Network Credit Scoring Models. *Computers & Operations Research*, nr 27, s. 1131-1152.

Summary

In the paper were presented researches constructing of model to estimating financial distress of enterprises based on artificial neural nets. Parameters of neural nets were optimized by scaled conjugate gradient and resilient backpropagation algorithms. In proposed models financial distress was developed based on information from quarterly reports of food branch enterprises quoted on WGPW (Warsaw Stock Exchange)

Adres do korespondencji:

dr inż. Henryk Marjak
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
Wydział Ekonomiczny
Zakład Analizy Systemowej
ul. K. Janickiego 31
71-270 Szczecin
tel. (91) 449 68 99
e-mail: henryk.marjak@zut.edu.pl