

Urszula Gierałtowska

Uniwersytet Szczeciński

WYKORZYSTANIE FUNKCJI DYSKRYMINACYJNEJ JAKO NARZĘDZIA WSPOMAGAJĄCEGO WYBÓR SPÓŁEK DO PORTFELA PAPIERÓW WARTOŚCIOWYCH

Gwałtowny rozwój rynku kapitałowego i osiągnięcie przez niego coraz większej dojrzałości rodzi potrzebę nieustannego tworzenia i doskonalenia metod analizy fundamentalnej, wspomagających proces podejmowania decyzji. Rynek kapitałowy jest organizmem złożonym, który fascynuje wielu badaczy pragnących wyjaśnić jego fenomen. Wpływ czynników egzogenicznych oraz psychologicznych sprawia jednak, że formułowanie uniwersalnych i ponadczasowych wniosków o jego funkcjonowaniu nie znajduje odzwierciedlenia w praktyce.

Rynek kapitałowy pozwala na przetworzenie informacji o sytuacji ekonomiczno-finansowej spółek w sygnały dotyczące cen poszczególnych walorów. Prawidłowe i wczesne rozpoznanie takich sygnałów pozwala inwestorowi na podejmowanie decyzji, które będą maksymalizować korzyści płynące z procesu inwestycyjnego. Najbardziej istotną kwestią jest dobór papierów wartościowych. Mimo że w doborze spółek do portfela nie ma metody doskonałej, coraz częściej zauważyć można, że wielu analityków próbuje adaptować i wdrażać metody statystyczno-ekonometryczne, które, z jednej strony, pozwalają na hierarchizację spółek giełdowych, a z drugiej – pośrednio oceniają dynamikę zmian rynku kapitałowego.

Ocena kondycji poszczególnych spółek notowanych na giełdzie jest jednym z podstawowych zagadnień, leżących w sferze zainteresowań zarówno analityków, jak i samych inwestorów, ponieważ możliwość niewypłacalności lub upadłości spółek ma poważne konsekwencje ekonomiczne i psychologiczne. W związku z tym, oprócz klasycznej analizy wskaźnikowej, coraz większe znaczenie zyskują metody wielowymiarowej analizy porównawczej, pozwalające na obiektywną ocenę kondycji ekonomiczno-finansowej oraz klasyfikację spółek według określonych wskaźników finansowych. Poszukuje się również nowszych narzędzi analitycznych, które

wraz z rozwojem technik komputerowych będą umożliwiać połączenie tradycyjnych teorii ekonomicznych z nowoczesnymi osiągnięciami statystyki i ekonometrii oraz staną się ogólnie dostępne i zrozumiałe dla większej liczby inwestorów.

Jest wiele metod klasyfikacji. W pracy zaproponowano wykorzystanie analizy dyskryminacyjnej. Funkcja dyskryminacyjna, nieco zapomniana przez badaczy, od kilku lat przeżywa renesans. Stosuje się ją m.in. w badaniach pedagogicznych, medycznych oraz coraz częściej w badaniach ekonomicznych, w tym również na rynku kapitałowym do klasyfikowania spółek giełdowych. Niewątpliwą wadą metody dyskryminacji jest to, że wymaga ona spełnienia dwóch podstawowych założeń dotyczących wielowymiarowego rozkładu normalnego zmiennych diagnostycznych oraz równości macierzy wariancji i kowariancji w poszczególnych klasach. W praktyce spełnienie tych założeń jest dużym ograniczeniem klasycznych metod dyskryminacji, ponieważ w znacznym stopniu może rzutować na jakość uzyskanych klasyfikacji. Z przeprowadzonych badań wynika jednak, że funkcja dyskryminacyjna, mimo niespełnienia omawianych warunków, może być dobrym klasyfikatorem, ponieważ jest odporna na te założenia¹.

Rozwiązaniem problemu może być również wykorzystanie metod, których stosowanie nie jest obwarowane tak surowymi założeniami. Do takich metod niewątpliwie zalicza się metody nieparametryczne, oparte na drzewach klasyfikacyjnych czy sieciach neuronowych, pozwalające na budowę modelu bez znajomości postaci rozkładów cech oraz związków między nimi.

Proces budowy drzewa klasyfikacyjnego odbywa się poprzez rekurencyjny podział zbioru obiektów na podzbiory, aż do uzyskania ich homogeniczności ze względu na przynależność do wyodrębnionych klas. Celem jest utworzenie drzewa o jak najmniejszej liczbie węzłów, aby otrzymać proste reguły klasyfikacyjne. Hierarchiczna natura oraz elastyczność drzew klasyfikacyjnych sprawiają, że są one wykorzystywane w naukach stosowanych takich, jak: medycyna, nauki komputerowe, psychologia, ponieważ za pomocą czytelnego i prostego obrazu graficznego ułatwiają interpretację skomplikowanych zjawisk wielowymiarowych².

W ostatnich latach coraz gwałtowniej wzrasta zainteresowanie – opartymi na wzorze mózgu człowieka – sztucznymi sieciami neuronowymi, zbudowanymi z neuronów, przetwarzającymi informacje w sposób równoległy. Sieci neuronowe stosowane są do rozwiązywania różnych problemów badawczych związanych z klasyfikacją oraz predykcją w różnorodnych dziedzinach takich, jak: medycyna, fizyka, chemia, inżynieria oraz ekonomia i zarządzanie³. Istotną cechą sieci jest umiejętność

¹ O zagadnieniach dyskryminacji traktują prace: [Jajuga 1990; Kolonko 1980] i inne.

² W literaturze polskiej i światowej można znaleźć wiele prac, które wskazują na możliwości aplikacyjne drzew klasyfikacyjnych w badaniach naukowych. Na szczególną uwagę zasługują prace E. Gatnara [2001].

³ Problematyka sztucznych sieci neuronowych nie jest tematem artykułu; jest szeroko omawiana w literaturze, np. w pracach W.S. McCullocha, W. Pittsa, E.M. Azoffa, C.M. Bishopa, R. Tadeusiewicza, J. Korbicza i innych. Krótki przegląd prac przykładowego zastosowania sieci neuronowych można znaleźć m.in. w pracach D. Witkowskiej [2002] oraz M. Lasek [2002].

abstrahowania, sprawiająca, że sieć neuronowa może uogólniać wiedzę zdobytą w procesie trenowania. Sztuczne sieci neuronowe szybko i efektywnie przetwarzają ogromne ilości danych, często chaotycznych, niekompletnych, a nawet sprzecznych, i funkcjonują nawet w momencie uszkodzenia lub zerwania połączenia.

W literaturze do badania kondycji oraz niewypłacalności przedsiębiorstw, oprócz przytoczonych metod, proponuje się również wykorzystanie zbliżonych do sieci neuronowych algorytmów genetycznych (cehuje je uniwersalność, elastyczność i skuteczność oraz duża odporność na zakłócenia występujące w danych wejściowych), modeli logitowych oraz łańcuchów Markowa, jak również rozmyte metody analizy skupień k -średnich.

Przykład empiryczny

W pracy zaprezentowane zostaną klasyczne i nieklasyczne metody klasyfikacji spółek giełdowych notowanych na Giełdzie Papierów Wartościowych w Warszawie, które w rozumieniu statystycznym stanowią zbiór obiektów wielowymiarowych, opisanych przez zmienne diagnostyczne o charakterze ogólnoeconomicznym, finansowym czy też organizacyjnym. Analizę ograniczono do spółek notowanych w systemie ciągłym i podwójnego fixingu w latach 2001-2003 (w ujęciu rocznym). Wyodrębniono zbiór 149 spółek giełdowych, dla których uzyskano pełen obraz wskaźników niezbędnych do przeprowadzenia dalszych analiz⁴.

Spośród wielu metod porządkowania liniowego wybrano dwie i w niniejszej pracy do oceny kondycji finansowej spółek zaproponowano wykorzystanie poziomu miar syntetycznych: taksonomicznej miary atrakcyjności inwestycji (TMAI)⁵ oraz uogólnionej miary odległości (GDM)⁶. Obie miary są unormowane w przedziale $[0; 1]$ – im wartość miary TMAI liczona dla i -tego obiektu jest bliższa jedności, tym obiekt w mniejszym stopniu różni się od wzorca rozwoju (w przypadku miary GDM sytuacja jest odwrotna). Informacja ta pozwoliła na budowę kryterium dyskryminacji spółek giełdowych. Inwestor wyodrębnił zatem grupę przedsiębiorstw:

- o dobrej kondycji finansowej, jeżeli $TMAI_i \geq M_{TMAI}$ ($GDM_i \leq M_{GDM}$),
- o złej kondycji finansowej, jeżeli $TMAI_i < M_{TMAI}$ ($GDM_i > M_{GDM}$).

Jako zmienne diagnostyczne, które posłużyły do oceny siły fundamentalnej spółek, w obu przypadkach zaproponowano wykorzystanie takich wskaźników, jak: płynność bieżąca, stopa zadłużenia, ROA, rotacja zapasów, rotacja należności oraz rotacja aktywów. Przy budowie miar syntetycznych przyjęto analogiczny spo-

⁴ Z badania wyeliminowano obiekty, dla których zanotowano luki w danych oraz banki, ponieważ zestaw wskaźników ekonomiczno-finansowych spółek sektorów niefinansowych nie jest tożsamy z zestawem wskaźników wykorzystywanych w ocenie standingu finansowego banków.

⁵ Po raz pierwszy miarę syntetyczną do klasyfikacji spółek ze względu na ich stopień atrakcyjności dla inwestora wykorzystał W. Tarczyński [1994, s. 275 i nast.].

⁶ Wykorzystanie uogólnionej miary odległości do klasyfikacji obiektów przedstawiono w pracy [Walesiak 2002].

sób ustalania wag – zmienne były jednakowo ważne z punktu widzenia danego problemu (tzn. miary bez wag) lub były ważone współczynnikiem zmienności (tzn. miary z wagami).

Po przyjęciu za kryterium dyskryminacji kondycji spółek giełdowych, mierzonych poziomem miary syntetycznej, dokonano estymacji parametrów funkcji dyskryminacyjnych, wykorzystując standardową i krokową analizę dyskryminacyjną. Jako zmienne diagnostyczne zaproponowano: stopę zwrotu (X_1), ryzyko (X_2), współczynnik beta (X_3), ROE (X_4), rotację zobowiązań (X_5), zysk (stratę) na 1 akcję (X_6), wartość księgową na 1 akcję (X_7). Zaproponowany subiektywny dobór wskaźników może znaleźć zarówno zwolenników, jak i przeciwników. Jednak mnogość wskaźników generowanych przez analizę wskaźnikową sprawia, że konieczne jest wyodrębnienie mierników, pozwalających na maksymalizację zawartości informacyjnej i kluczowych w podejmowaniu decyzji. Brak metody naukowej, pozwalającej na wybór najlepszych czy najskuteczniejszych wskaźników, sprawia, że dobór ten może być oceniony jedynie *ex post* pod kątem uzyskanych efektów inwestowania.

Wstępne wyniki klasyfikacji spółek giełdowych przy wykorzystaniu miar TMAI oraz GDM były zbliżone (tab. 1). Znalazło to swoje odbicie w oszacowaniach parametrów. W przypadku wykorzystania standardowej analizy dyskryminacyjnej można zauważyć niewielkie różnice w wartościach oszacowanych parametrów, zgodny jest zaś rząd szacowanych wielkości. Nieco odmienna sytuacja wystąpiła, gdy do szacowania wykorzystano krokową postępującą analizę dyskryminacyjną. Niewielkie różnice w optymalnym wektorze zmiennych objaśniających można zauważyć jedynie w przypadku wykorzystania ważonych miar syntetycznych jako kryterium poziomu. W przypadku wykorzystania nieważonych miar syntetycznych do ustalenia wstępnego podziału spółek inne są zarówno wartości wskaźników, jak i optymalny zestaw zmiennych, mających największą moc dyskryminacyjną. Ogólny błąd klasyfikacji mieścił się w przedziale 13,4-25,5% dla kryterium uwzględniającego wagi, a 20,8-37,6% dla kryterium nieuwzględniającego wag. Należy również zauważyć, że zastosowanie uogólnionej miary odległości GDM (w porównaniu z TMAI) do wstępnego podziału spółek giełdowych pozwoliło na osiągnięcie większej dokładności w przypadku wykorzystania krokowej analizy dyskryminacyjnej, a gorszej przy analizie standardowej.

W celu porównania przeprowadzono klasyfikację spółek na podstawie metod nieklasycznych, wykorzystując w tym celu pakiet statystyczny STATISTICA 6.0 (moduł drzewa klasyfikacyjne oraz sieci neuronowe). W drzewach klasyfikacyjnych eksperymenty badawcze przeprowadzono z wykorzystaniem dyskryminacyjnych podziałów jednowymiarowych oraz metody CART (metoda dyskrymina-

Tabela 1. Podstawowe statystyki opisujące moc dyskryminacyjną funkcji dyskryminacyjnych

	2001						2002						2003					
	TMAL bez wag			TMAL z wagami			TMAL bez wag			TMAL z wagami			TMAL bez wag			TMAL z wagami		
	S	K		S	K		S	K		S	K		S	K		S	K	
λ_w	0,84	0,8426		0,6908	0,6966		0,7085	0,7156		0,6603	0,6645		0,762	0,7734		0,6816	0,6952	
F	3,8351	5,3409		9,0175	12,4576		8,2859	11,3655		10,3637	14,4411		6,2909	14,1632		9,4081	21,1854	
χ^2	25,0098	24,7405		53,0896	52,2469		49,443	48,3511		59,5655	59,0642		39,0021	37,3919		54,999	52,8655	
X_1	-1,2857			28,262	21,7479		22,2633	21,8639		29,9085	32,7827		17,2711			21,4434		
X_2	-15,6375	-13,7194		5,5971			-3,2488			1,2776			-7,618			-7,9345		
X_3	0,2885			0,1512			0,2741			0,2758	0,2625		0,0605			-0,0269		
X_4	0,355	0,3867		0,2586	0,2798		0,0884	0,0933		0,0692			0,5433	0,6381		1,0988	1,141	
X_5	0,0012	0,0013		0,0016	0,0017		-0,0004	-0,0004		0,0006	0,0006		0,0001			0,0001		
X_6	0,1271	0,1316		0,2043	0,1917		0,1636	0,1685		0,2051	0,2106		0,1144	0,1324		0,2541	0,2821	
X_7	0,0108	0,0108		0,0145	0,0141		0,0103	0,012		0,0115	0,0112		0,0323	0,0355		-0,0183	-0,0161	
Stała	0,6246	0,6396		-0,6781	-0,2775		0,3006	0,17		-0,2331	-0,1566		-0,327	-0,7021		0,3518	-0,0534	
Dokładność	43 32 43	32 58 17	58 17 51	24 51 24	52 23 55	20 58 17	59 16 61	14 69 6										
E	18 56 19	55 15 59	10 64 9	65 7 67	15 59 15	26 48 26	48 15 59	15 59 15		25,5	23,49		28,86	28,19		19,46	14,09	
	33,56	34,23		21,48	18,12		22,15	20,81		25,5	23,49		28,86	28,19		19,46	14,09	
	GDM bez wag			GDM z wagami			GDM bez wag			GDM z wagami			GDM bez wag			GDM z wagami		
	S	K		S	K		S	K		S	K		S	K		S	K	
λ_w	0,8988	0,9126		0,6908	0,6966		0,704	0,7139		0,6618	0,6673		0,7794	0,789		0,6789	0,6917	
F	2,2674	4,6271		9,0175	12,4576		8,4704	19,367		10,2928	17,9482		5,7	9,6297		9,528	21,5391	
χ^2	15,307	13,3023		53,0896	52,2469		50,3717	49,029		59,2317	58,6533		35,7578	34,3707		55,5799	53,624	
X_1	-7,1961			28,262	21,7479		36,5263	37,2354		31,9012	35,2613		13,9118			19,9415		
X_2	13,6826			5,5971			-5,1828	-6,3857		1,0913			-6,1884			-7,8529		
X_3	-0,5636			0,1512			0,1961			0,2837	0,2744		0,3537	0,3982		-0,0978		
X_4	0,4435	0,4191		0,2586	0,2798		0,0549			0,0674			0,3362			1,1016	1,1335	
X_5	0,0026	0,0027		0,0016	0,0017		0,0001			0,0006	0,0006		0,0002	0,0002		0,0001		
X_6	0,0821			0,2043	0,1917		0,1542	0,1764		0,2145	0,2312		0,1602	0,1963		0,2588	0,2859	
X_7	0,0216	0,0228		0,0145	0,0141		0,0103			0,0059			0,0244	0,0262		-0,0193	-0,0171	
Stała	-1,395	-0,8059		-0,6781	-0,2775		0,3968	0,7412		-0,1226	0,0368		-0,5211	-0,9215		0,425	-0,0405	
Dokładność	58 17 61	14 58 17	58 17 50	25 51 24	51 24 51	24 52 23	55 20 58	17 62 13										
E	39 35 42	32 15 59	10 64 11	63 11 63	12 62 25	49 26 48	15 59 14	60 13 69 6		23,49	23,49		30,2	28,86		18,79	13,42	
	37,58	37,58		21,48	18,12		24,16	23,49		23,49	23,49		30,2	28,86		18,79	13,42	

W polu dokładność – wartości $\frac{a}{b}$ oznaczają: a/d – liczbę poprawnie zaklasyfikowanych obiektów niespełniających/spełniających kryterium, b/c – liczbę

niepoprawnie zaklasyfikowanych obiektów niespełniających/spełniających kryterium.

E (błąd klasyfikacji) oznacza odsetek błędnie zaklasyfikowanych obiektów.

Źródło: obliczenia własne.

cyjnych podziałów jednowymiarowych dla predyktorów nominalnych i porządkowych), w przypadku sieci neuronowych wykorzystano najpopularniejszą architekturę sieciową opartą na koncepcji perceptronu wielowarstwowego (MLP – *multi-layer perceptron*). Dokładność klasyfikacji spółek giełdowych (mierzoną odsetkiem błędnie zaklasyfikowanych obiektów) zamieszczono w tab. 2.

Tabela 2. Ocena błędnej klasyfikacji spółek giełdowych przy wykorzystaniu klasycznej analizy dyskryminacyjnej, drzew klasyfikacyjnych i sieci neuronowych

	FD S	FD K	D1	D2	S
2001					
TMAI bw	33,56	34,23	36,24	26,17	34,9
TMAI w	21,48	18,12	16,78	18,12	22,15
GDM bw	37,58	37,58	38,25	25,5	39,6
GDM w	21,48	18,12	16,78	18,12	22,15
2002					
TMAI bw	22,15	20,81	16,78	17,45	36,91
TMAI w	25,5	23,49	16,11	14,76	26,84
GDM bw	24,16	23,49	28,86	22,82	28,19
GDM w	23,49	23,49	14,76	14,09	20,13
2003					
TMAI bw	28,86	28,19	27,52	26,84	26,84
TMAI w	19,46	14,09	16,78	12,75	17,44
GDM bw	30,2	28,86	32,89	15,44	29,53
GDM w	18,79	13,42	10,74	12,75	10,74

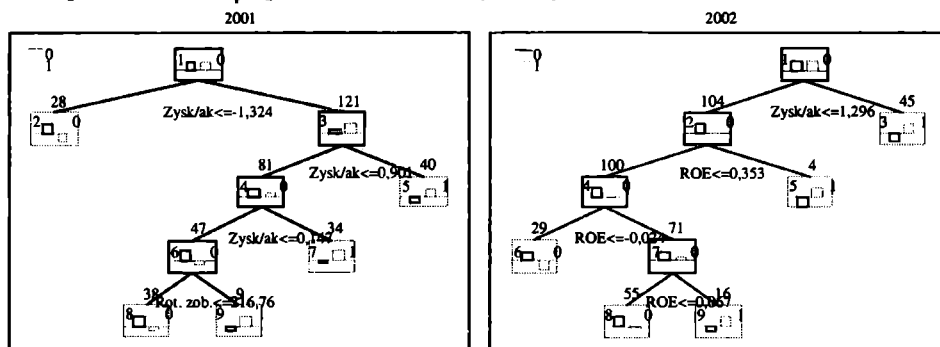
Metody: FD S – standardowa analiza dyskryminacyjna, FD K – krokowa analiza dyskryminacyjna, D1 – dyskryminacyjne podziały jednowymiarowe, D2 – metoda CART, S – sieć neuronowa – perceptron trójwarstwowy.

Źródło: obliczenia własne.

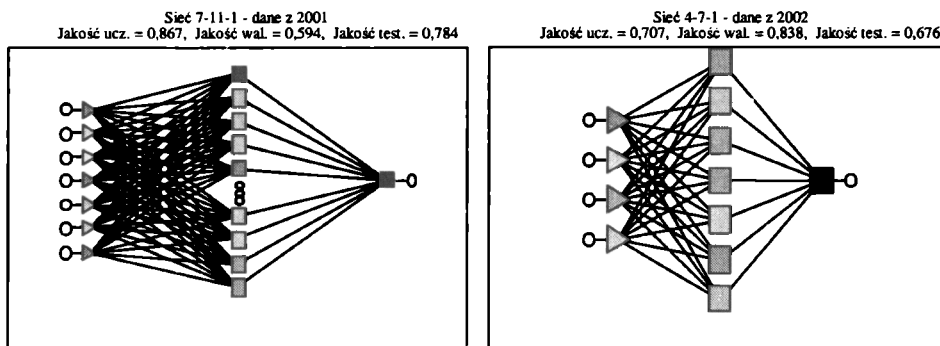
W analizie wyników zawartych w tab. 2 należy zauważyć, że nieznacznie lepsze wyniki klasyfikacji spółek giełdowych uzyskano z wykorzystaniem drzew klasyfikacyjnych, przy czym znacznie wyższą jakością charakteryzowała się klasyfikacja z wykorzystaniem metody CART, gwarantującej znalezienie podziału, który stanowi najlepszą klasyfikację próby uczącej. Wykorzystując sztuczne sieci neuronowe, można zauważyć, iż w większości przypadków gorzej klasyfikują spółki w porównaniu z klasyfikacją z wykorzystaniem funkcji dyskryminacyjnej czy drzew klasyfikacyjnych. Prezentowane sieci neuronowe dużo lepiej radzą sobie z klasyfikacją obiektów w momencie, kiedy na wyjściu dysponują ważonymi miarami syntetycznymi opisującymi kondycję. Na rysunkach 1-2 przedstawiono przykładową graficzną prezentację wyników klasyfikacji (TMAI z wagami) za pomocą drzew klasyfikacyjnych oraz sieci neuronowych.

Otrzymane wyniki pozwalają twierdzić, że zarówno drzewa klasyfikacyjne, jak i sieci neuronowe mogą być wykorzystywane jako alternatywa dla klasycznych metod klasyfikacji, zwłaszcza jeśli wiadomo, że nie są spełnione założenia analizy dyskryminacyjnej (dotyczy to przede wszystkim rozkładu wskaźników finansowo-ekonomicznych). Należy jednocześnie zauważyć, że wiele ze wskaźników eko-

nomiczno-finansowych nie wpływa liniowo na klasę obiektów, dlatego podział zbioru z wykorzystaniem liniowej funkcji dyskryminacyjnej często bywa przypadkowy. Drzewa klasyfikacyjne (a zwłaszcza metoda CART) nie ograniczają się do jednej liniowej kombinacji zmiennych; umożliwiają znalezienie takich punktów, które pozwolą na najlepszą możliwą klasyfikację obiektów.



Rys. 1. Drzewa klasyfikacyjne (dyskryminacyjne podziały jednowymiarowe – D1) dla kryterium TMAI z wagami
Źródło: opracowanie własne.



Rys. 2. Przykładowa struktura sieci MLP dotycząca danych rocznych – kryterium TMAI z wagami
Źródło: opracowanie własne.

Konstrukcja zaprezentowanych metod opiera się na danych historycznych. Mogą być one wykorzystane do klasyfikowania spółek giełdowych do grup wyodrębnionych w przyszłości. Z celu zbadania stopnia trafności klasyfikacji i użyteczności funkcji dyskryminacyjnej oraz metod nieklasycznych można przetestować otrzymane reguły klasyfikacyjne na próbie walidacyjnej. W przypadku funkcji dyskryminacyjnej do rozstrzygnięcia, do której grupy należą poszczególne obiekty badawcze wykorzystano funkcje klasyfikacyjne. Reguła klasyfikacyjna jednoznacznie określa, że obiekt zaklasyfikowany zostaje do grupy, dla której wartość funkcji przyjmuje najwyższą wartość (przy założeniu że prawdopodobieństwa *a priori* nie różnią się istotnie między sobą). Na podstawie reguł klasyfikacyjnych z

2001 r. dokonano klasyfikacji obiektów w 2002 i 2003 r., na podstawie reguł z 2002 r. zaprognozowano zaś pozycję spółek w 2003 r. Błędy klasyfikacji prognoz przedstawiono w tab. 3.

W większości analizowanych przypadków jakość przeprowadzonej klasyfikacji nowych obiektów była najwyższa z zastosowaniem drzew klasyfikacyjnych, a najniższa – z użyciem sieci neuronowych. Najbardziej istotne różnice zaobserwować można w przypadku wykorzystania do prognoz reguł klasyfikacyjnych, wyznaczonych na podstawie danych z 2002 r. Można jednocześnie zauważyć, iż niezależnie od przyjętej procedury klasyfikacji spółek i okresu badawczego, dużo lepsze wyniki klasyfikacji uzyskano wówczas, gdy do oceny kondycji spółek ważonych wykorzystano zmienne syntetyczne (GDM oraz TMAI).

Tabela 3. Ocena dokładności prognoz klasyfikacji spółek giełdowych

	FD S	FD K	D1	D2	S
2001 → prognoza na 2002					
TMAI bw	30,20	29,53	22,82	24,16	33,56
TMAI w	20,81	20,14	16,78	16,11	24,16
GDM bw	40,94	45,75	62,42	53,02	55,70
GDM w	20,81	20,14	16,78	16,11	24,16
2001 → prognoza na 2003					
TMAI bw	37,58	36,91	34,23	29,53	65,77
TMAI w	38,93	38,26	27,52	23,49	63,76
GDM bw	42,95	39,60	60,40	48,99	51,01
GDM w	38,93	38,26	27,52	23,49	63,76
2002 → prognoza na 2003					
TMAI bw	42,95	43,63	23,49	30,87	68,46
TMAI w	40,27	40,27	22,82	24,83	54,36
GDM bw	46,31	44,30	40,27	33,56	67,78
GDM w	40,27	40,27	22,15	26,84	67,78

Źródło: obliczenia własne.

Zaskakują słabe wyniki klasyfikacji pozycji spółek przy wykorzystaniu sieci neuronowych. Należy jednak zwrócić uwagę, że stosowanie sieci neuronowych do klasyfikacji obiektów wymaga odpowiedniego przygotowania zbioru uczącego i testowego. Jedynie odpowiednio wytrenowana i przetestowana sieć może być wykorzystywana do oceny kolejnych obiektów, które nie były zastosowane ani w procedurze uczenia, ani testowania.

Wyniki przeprowadzonych badań pozwalają stwierdzić, że funkcja dyskryminacyjna może być wykorzystywana do dyskryminacji spółek giełdowych, ponieważ zarówno wyniki klasyfikacji obiektów, jak i prognozy klasyfikacji w wielu przypadkach nie odbiegają zasadniczo od wyników uzyskanych z użyciem metod nieparametrycznych. Zarówno drzewa klasyfikacyjne, jak i sieci neuronowe mają prostą i łatwą do interpretacji strukturę, jednak z punktu widzenia inwestora, w poszukiwaniu dobrych modeli predykcyjnych, konieczne jest dokładne zrozumienie natury badanych zależności. Zbyt rozbudowane procedury sprawiają, że ich interpre-

tacja jest dużo bardziej skomplikowana niż jednej liniowej funkcji dyskryminacyjnej. Należy również dodać, że funkcja dyskryminacyjna ma tę przewagę nad metodami nieklasycznymi, że nie tylko klasyfikuje obiekty do wyodrębnionych klas, ale pozwala również na ich uporządkowanie, co z punktu widzenia procesu inwestowania jest bardziej istotną kwestią⁷.

Literatura

- Gatnar E., *Nieparametryczna metoda dyskryminacji i regresji*, PWN, Warszawa 2001.
- Gierałtowska U., *Możliwości wykorzystania funkcji dyskryminacyjnej na polskim rynku kapitałowym*, [w:] *Rynek kapitałowy. Skuteczne inwestowanie*, Zeszyty Naukowe Uniwersytetu Szczecińskiego nr 389, cz. II, Szczecin 2004.
- Jajuga K., *Statystyczna teoria rozpoznawania obrazów*, PWN, Warszawa 1990.
- Kolonko J., *Analiza dyskryminacyjna i jej zastosowania w ekonomii*, PWN, Warszawa 1980.
- Lasek M., *Data Mining. Zastosowania w analizach i ocenach klientów bankowych*, Biblioteka Menedżera i Bankowca, Warszawa 2002.
- Tarczyński W., *Taksonomiczna miara atrakcyjności inwestycji w papiery wartościowe*, Przegląd Statystyczny, R. XLI, z. 3, Warszawa 1994.
- Walesiak M., *Uogólniona miara odległości w statystycznej analizie wielowymiarowej*, AE, Wrocław 2002.
- Witkowska D., *Sztuczne sieci neuronowe i metody statystyczne. Wybrane zagadnienia finansowe. Studia ekonomiczne*, Wydawnictwo C.H. Beck, Warszawa 2002.

USING DISCRIMINANT FUNCTION AS AN INSTRUMENT OF PORTFOLIO SELECTION IN THE CAPITAL MARKET

Summary

In the article the author carried out the usage of discriminant function with a agglomerative measure as a criterion of discrimination in order to evaluate economic and financial condition of a company. The companies at the Warsaw Stock Exchange were classified on the basis of economic and financial measures in the years 2001–2003. The results of the classification were acceptable and they were used to predict to which of the two groups a company will belong. The results of this evaluation and analysis were compared with the results of nonclassical methods. The structural assumptions of the function were not fulfilled, but the function was a good identifier all the same.

⁷ Szczegółowe wyniki empiryczne przedstawiono w pracy [Gierałtowska 2004].