

Łukasz Mach

Politechnika Opolska

Zastosowanie analizy dyskryminacyjnej i sieci neuronowych do oceny możliwości sprzedaży nieruchomości mieszkaniowych

Streszczenie. W artykule przedstawiono zastosowanie analizy dyskryminacyjnej oraz narzędzi sztucznych sieci neuronowych do oceny możliwości sprzedaży nieruchomości mieszkaniowych. Autor podjął próbę zastosowania oraz porównania metody zaliczanej do kategorii statystycznych metod wielowymiarowych (analiza dyskryminacyjna) z metodą bazującą na zastosowaniu sztucznej inteligencji (sieci neuronowe). Celem zastosowania wybranych metod było sprawdzenie ich zdolności dyskryminacyjnych w procesie klasyfikacji nieruchomości mieszkaniowych. Klasyfikacji do dwóch rozłącznych grup, tj. grupy mieszkań rokujących na sprzedaż w określonym czasie oraz grupy mieszkań niedających szans na sprzedaż. Poprawnie przygotowane modele analityczne bazujące na analizie dyskryminacyjnej oraz sieciach neuronowych mogą wpłynąć na poprawę konkurencyjności przedsiębiorstw działających w obszarze rynku nieruchomości, zwłaszcza w czasie globalnego kryzysu gospodarczego.

Słowa kluczowe: analiza dyskryminacyjna, sieci neuronowe, rynek nieruchomości mieszkaniowych

Wstęp

Poprawnie przygotowany oraz przeprowadzony proces podejmowania decyzji jest istotnym czynnikiem oddziałującym na wzrost konkurencyjności przedsiębiorstw. Dobrze przygotowany proces decyzyjny pozwala zminimalizować

ryzyko źle podjętych decyzji, pozwala także wzmocnić pozycję konkurencyjną przedsiębiorstw¹.

Artykuł jest kontynuacją badań w obszarze rynku nieruchomości mieszkaniowych, w których zostały opracowane modele decyzyjne z wykorzystaniem regresji wielorakiej² oraz regresji logistycznej³.

Nadrzędny cel artykułu skupia się na przedstawieniu potencjalnych możliwości klasyfikacji wybranych nieruchomości mieszkaniowych do jednej z dwóch grup. Do grupy, gdzie sprzedaż mieszkania nastąpi wcześniej niż średni czas oczekiwania mieszkań na sprzedaż oraz do grupy, w której znalazły się mieszkania o czasie sprzedaży powyżej średniego czasu oczekiwania mieszkań na sprzedaż.

Artykuł składa się z trzech podstawowych części. W pierwszej przedstawiono założenia formalne korzystania z metod analizy dyskryminacyjnej i sztucznych sieci neuronowych. W drugiej części został opisany proces badawczy, w którym opracowano oraz sprawdzono użyteczność opracowanych modeli klasyfikacyjnych. W końcowej, trzeciej, części artykułu przedstawiono wnioski z przeprowadzonych badań oraz dokonano ich interpretacji.

1. Realizacja procesu badawczego – założenia formalne

Realizacja procesu badawczego została oparta na zastosowaniu dwóch narzędzi analizy danych. Zastosowanie analizy dyskryminacyjnej miało na celu określenie oraz zbadanie różnic między dwiema grupami nieruchomości, z jednoczesnym uwzględnieniem zmiennych je różnicujących. Rozpatrywana analiza została odniesiona do kilku wzajemnie powiązanych procedur statystycznych⁴:

- procedur opisujących i interpretujących międzygrupowe zróżnicowanie,
- procedur klasyfikacji przypadków.

¹ Patrz.: M. Adamska, *Diagnoza czynników warunkujących rozwój małych i średnich przedsiębiorstw*, w: *Od naukowej inspiracji do innowacji w przedsiębiorstwie. Praktyczna aplikacja wiedzy Asystentów innowacji*, red. M. Szewczuk-Stępień, M. Adamska, Instytut Trwałego Rozwoju, Opole 2013, s. 36-37.

² Ł. Mach, *Budowa praktycznego modelu regresji opisującej zależności występujące na rynku nieruchomości mieszkaniowych*, red. S. Forlicz, „Zeszyty Wyższej Szkoły Bankowej we Wrocławiu” 2011, nr 20, s. 291-302.

³ Ł. Mach, *Modelowanie logitowe narzędziem wspomagającym podejmowanie decyzji na rynku nieruchomości mieszkaniowych*, w: *Prognozowanie w zarządzaniu firmą*, red. P. Dittmann, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu” 2010, s. 106-113.

⁴ A. Stanisławski, *Przystępny kurs statystyki z zastosowaniem Statistica PL na przykładach z medycyny*, t. 3, *Analizy wielowymiarowe*, Statsoft, Kraków 2007, s. 51.

W pierwszym kroku zostały wyznaczone kanoniczne funkcje dyskryminacyjne.

Z uwagi na powszechnie stosowaną postać liniową parametryzowane funkcje wyrażono zależnością⁵:

$$D_{kj} = \beta_0 + \beta_1 x_{1kj} + \dots + \beta_p x_{pkj}, \quad (1)$$

gdzie: p – liczba zmiennych dyskryminacyjnych, D_{kj} – wartości kanonicznej dyskryminacyjnej funkcji dla k -tego przypadku w j -tej grupie, $k = 1, \dots, n$ (n – liczebność próby) oraz $j = 1, \dots, g$ (g – liczba grup), $i = 1, \dots, p$, x_{ikj} – wartość i -tej zmiennej dyskryminacyjnej dla k -tego przypadku w j -tej grupie; β_i – współczynniki kanonicznej funkcji dyskryminacyjnej wyznaczone na podstawie wartości tej funkcji.

Wyznaczając współczynniki funkcji dyskryminacyjnej wyznaczono maksymalny stosunek zmienności międzygrupowej zmiennych wejściowych do ich zmienności wewnątrzgrupowej:

$$(M - \lambda_j W) a_j = 0, \quad j = 1, 2, \dots, s, \quad (2)$$

gdzie: M – macierz międzygrupowej sumy kwadratów

$$M = \sum_{r=1}^z n_r (\bar{z}_r - \bar{z})^T (\bar{z}_r - \bar{z}), \quad (3)$$

W – macierz wewnątrzgrupowej sumy kwadratów

$$W = \sum_{r=1}^z \sum_{i=1}^{n_r} (z_{ri} - \bar{z}_r)^T (z_{ri} - \bar{z}_r), \quad (4)$$

gdzie: $z_{ri} = [z_{rij}]$ – wektor standaryzowanych zmiennych w r -tej grupie obiektów, przy czym z_{rij} jest wartością j -tą zmiennej w i -tym obiekcie należącym do r -tej grupy; $\bar{z}_r = [\bar{z}_{rj}]$ – wektor średnich wartości standaryzowanych zmiennych w r -tej grupie, przy czym \bar{z}_{rj} jest średnią wartością j -tej zmiennej w r -tej grupie; $\bar{z} = [\bar{z}_j]$ – wektor średnich wartości standaryzowanych zmiennych, przy czym \bar{z}_j jest średnią wartością j -tej grupy.

Układ ma niezerowe rozwiązanie, gdy spełniony jest warunek:

$$|M - \lambda W| = 0. \quad (5)$$

Prowadzi to do wyznaczenia szeregu współczynników tzw. wartości własnych $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_w$, którym przypisane są wektory własne $\beta = (\beta_0, \beta_1, \dots, \beta_p)$.

⁵ A.D. Aczel, *Statystyka w zarządzaniu*, WN PWN, Warszawa, 2000, s. 882.

Wektory własne stanowią wyliczane współczynniki kanonicznej funkcji dyskryminacyjnej⁶.

Przyjęto również założenie, że jeśli analizowane zmienne nie są w postaci standaryzowanej, w celach interpretacyjnych wyników poddaje się je standaryzacji:

$$\hat{\beta}_i = \beta_i \sqrt{n - g}, \quad (6)$$

$$\hat{\beta}_0 = -\sum_{i=1}^p \hat{\beta}_i \bar{x}_i, \quad (7)$$

gdzie: n – liczebność próby, g – liczba grup, \bar{x}_i – średnia wartość i -tej zmiennej dyskryminacyjnej dla wszystkich przypadków, β_i – surowe współczynniki kanonicznej funkcji dyskryminacyjnej $i = 1, \dots, p$, $\hat{\beta}_i$ – standaryzowane współczynniki funkcji dyskryminacyjnej $i = 1, \dots, p$, p – liczba zmiennych dyskryminujących.

Współczynniki standaryzowane zostały wyznaczone w jednostkach odchylenia standardowego i wskazują siłę wpływu wybranej zmiennej dyskryminacyjnej na zróżnicowanie grupy. Im większa jest jego wartość, tym silniejszy jest wpływ dyskryminacyjny rozpatrywanej zmiennej.

Podsumowanie zastosowania opisanych powyżej metod statystycznych powinno się rozpocząć od zbadania istotności wyznaczonych funkcji dyskryminacyjnych, następnie należy ocenić współczynniki struktury czynnikowej.

Drugim etapem zastosowanym w analizie dyskryminacyjnej było przewidywanie klasyfikacji przypadków za pomocą funkcji klasyfikacyjnych wyrażonej zależnością:

$$K_i = c_{i0} + c_{i1}x_1 + c_{i2}x_2 + \dots + c_{ip}x_p, \quad (8)$$

gdzie c_{ij} , $i = 0, \dots, p$, to specjalne współczynniki zmiennych dyskryminacyjnych. Funkcje klasyfikacyjne zostały wykorzystane w celu stwierdzenia, do której grupy należy rozpatrywany przypadek (przypadek jest zaliczany do grupy, dla której K_i przyjmuje wartość największą).

Należy również zaznaczyć, iż kluczowym zagadnieniem w analizie dyskryminacyjnej jest wnioskowanie o:

- prawdopodobieństwie *a priori* dla polepszenia dokładności predykcji:

$$P(G_k / X) = \frac{p_k P(X / G_k)}{\sum_{i=1}^g p_i P(X / G_i)}, \quad (9)$$

⁶ T. Panek, *Statystyczne metody wielowymiarowej analizy porównawczej*, Wyd. Szkoły Głównej Handlowej w Warszawie, Warszawa, 2009, s. 279.

gdzie: g – liczba grup, p_i – prawdopodobieństwa *a priori*, $P(G_k/X)$ – prawdopodobieństwo warunkowe tego, iż dany przypadek należy do grupy k , uwarunkowane znajomością zmiennych dyskryminacyjnych, $P(X/G_k)$ – prawdopodobieństwo warunkowe otrzymania wektora zmiennych, jeśli wiemy, że przypadek należy do grupy k ;

– oczekiwanym prawdopodobieństwie błędnej klasyfikacji:

$$\sum_{i=1}^g p_i \sum P(k/i), \quad (10)$$

gdzie: $P(k/i)$ – prawdopodobieństwo błędnego zaklasyfikowania obiektu z i -tej grupy do grupy k -tej⁷.

Dokonując opisu założeń przyjętych dla użytych sieci neuronowych, należy na wstępie zaznaczyć, iż dla potrzeb dyskryminacyjnych zostały wybrane sieci neuronowe perceptronowe wielowarstwowe⁸. Topologia użytych sieci została zdeterminowana przez rozwiązywany problem dyskryminacyjny. Ponieważ z założenia działania sieci neuronowych warstwa wyjściowa generuje sygnały wyjściowe dla całej sieci, więc przyjęto założenie, że liczba elementów tej warstwy będzie odpowiadać liczbie zmiennych charakteryzujących wymiar rozwiązywanego zadania. Z kolei użycie warstw ukrytych znacząco zwiększyło możliwości odwzorowań, które sieć potrafiła realizować. Liczba wejść do każdej warstwy odpowiada liczbie neuronów, jaka znajduje się w poprzedniej warstwie. Uwzględniając powyższe założenia, wartość wyjściowa j -tego neuronu w l -tej (dla $l = 2, 3, \dots, L - 1$) warstwie ukrytej jest wyznaczana na podstawie wyrażenia⁹:

$$h_j^l = F_l \left(\sum_{k=0}^{K^{l-1}} w_{jk}^l h_k^{l-1} \right), \quad (11)$$

natomiast dla pierwszej warstwy ukrytej poszczególne elementy w niej zawarte zdefiniowane zostały przez wyrażenie¹⁰:

$$h_j^1 = F_1 \left(\sum_{i=0}^n w_{jk}^1 x_i \right). \quad (12)$$

⁷ A. Stanisławski, *Przystępny kurs statystyki z zastosowaniem Statistica PL na przykładach z medycyny*, t. 3, *Analizy wielowymiarowe*, Statsoft, Kraków 2007, s. 67.

⁸ S. Osowski, *Sieci neuronowe do przetwarzania informacji*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2000.

⁹ D. Witkowska, *Sztuczne sieci neuronowe i metody statystyczne*, Wyd. C.H. Beck, Warszawa 2002, s. 11.

¹⁰ Ibidem, s. 12.

Ponadto wartość j -tego neuronu w warstwie wyjściowej ($j = 1, 2, \dots, m$) wyznaczono z wykorzystaniem wyrażenia:

$$y_j = F_L \left(\sum_{k=0}^{K^{L-1}} w_{jk}^L h_k^{L-1} \right). \quad (13)$$

Wartości neuronów w ostatniej warstwie były jednocześnie wartościami wyjściowymi z sieci.

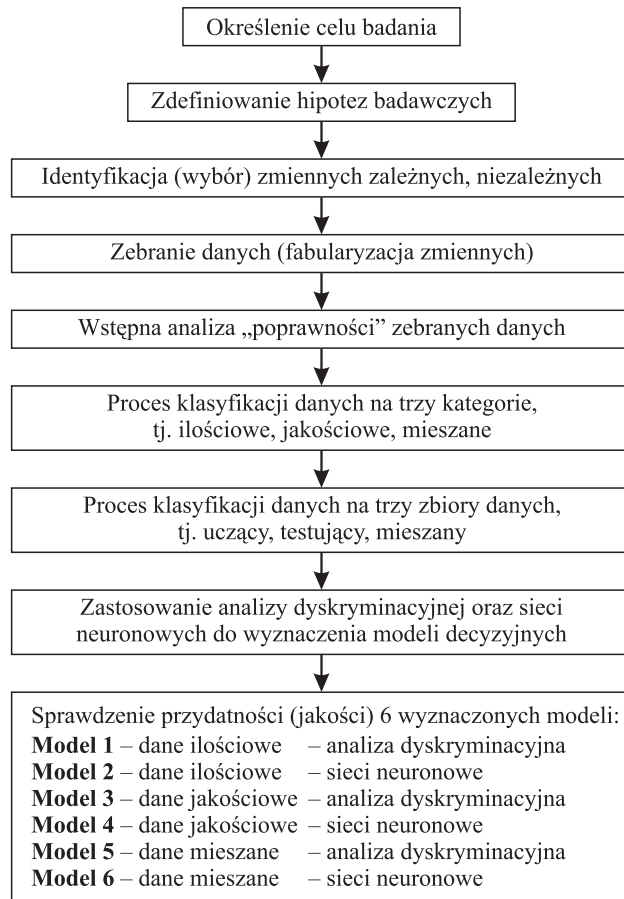
Przyjmując założenie, że określono $(L - 1)$ warstw ukrytych, w każdej po K^l elementów, przyjmujących oznaczenia: h_k^l – wartość wyjściowa k -tego neuronu w l -tej warstwie ukrytej, x_i – wartość wejścia i -tego neuronu w warstwie wejściowej, y_j – wartość wyjścia j -tego neuronu w warstwie wyjściowej, F_l – symbole funkcji aktywacji, $w_{jk}^l, w_{ji}^l, w_{jk}^L$ – wagi wyznaczone dla poszczególnych elementów we wszystkich warstwach.

Dodatkowe założenia przyjęte dla użytych sieci neuronowych to: podział zbioru danych na uczące i testujące w proporcji 80% do 20%, określenie minimalnej liczby neuronów w warstwach ukrytych na 4 neurony, określenie maksymalnej liczby neuronów w warstwie ukrytej na 14. Dobierając parametry sieci neuronowych w warstwie ukrytej przebadano funkcję aktywacji w postaci: funkcji liniowej, funkcji logistycznej, tangensu hiperbolicznego oraz funkcji wykładniczej. Optymalizując działanie sieci zastosowano funkcję błędu sum kwadratów.

2. Proces badawczy – zastosowanie analizy dyskryminacyjnej oraz sieci neuronowych w praktyce

Przeprowadzony proces badawczy bazował na zastosowaniu analizy dyskryminacyjnej oraz sieci neuronowych do oceny możliwości sprzedaży nieruchomości mieszkaniowych. Przeprowadzając badania autor wyznaczył sześć różnych modeli analitycznych. Dwa modele z wykorzystaniem wyłącznie danych ilościowych, dwa wykorzystujące wyłącznie dane jakościowe oraz dwa tzw. modele mieszane (tzn. z użyciem łącznie danych ilościowych oraz jakościowych). W modelach opartych na danych jakościowych, ilościowych oraz mieszanych zastosowano w pierwszej kolejności analizę dyskryminacyjną, następnie narzędzia sztucznych sieci neuronowych. W każdym modelu sprawdzono jego przydatność na zbiorze danych uczących i testujących. Algorytm (etapy) przeprowadzonych badań przedstawiono na rysunku 1.

W pierwszym etapie badawczym został określony cel badania, którym było zbudowanie modeli analitycznych, dyskryminujących wybraną nieruchomość do jednej z dwóch grup: tak zwanej grupy 0, gdzie czas sprzedaży mieszkania jest



Rys. 1. Algorytm przeprowadzonych badań

Źródło: opracowanie własne.

możliwy poniżej średniego czasu oczekiwania mieszkań na sprzedaż, oraz tzw. grupy 1, gdzie znalazły się mieszkania, których czas sprzedaży jest powyżej średniego czasu oczekiwania mieszkań na sprzedaż.

Badania empiryczne realizowane w artykule pozwolą m.in. odpowiedzieć na następujące hipotezy badawcze:

- zastosowanie analizy dyskryminacyjnej daje lepsze wyniki w modelowaniu rynku nieruchomości mieszkaniowych niż zastosowanie narzędzi sieci neuronowych,
- w analizie danych transakcyjnych dane ilościowe mają mniejsze znaczenie niż dane jakościowe.

Budowa wszystkich modeli bazowała na transakcjach sprzedaży nieruchomości mieszkaniowych dokonanych w jednym z miast wojewódzkich od I kw.

2010 r. do III kw. 2011 r. (dane transakcyjne). Przeprowadzając proces badawczy w pierwszej kolejności sprawdzono, czy występują obserwacje odstające. Proces eliminacji obserwacji odstających przeprowadzono w odniesieniu do każdej zmiennej mierzalnej. W tym celu na bazie wykresu ramka-wąsy oraz obliczonych miar położenia (mediana, kwartył1, kwartył3) odrzucono wszystkie obserwacje, które mogłyby być uznane za odstające (za kryterium klasyfikacji zmiennej jako odstającej przyjęto ± 3 IQR).

Następnie dokonano wyboru zmiennych objaśniających, mogących mieć istotny wpływ na rozpatrywane zjawisko oraz wstępnej ich analizy.

Do zbioru zmiennych wybrano następujące determinanty (X):

- cena za metr kwadratowy,
- powierzchnia mieszkania,
- liczba pokoi w mieszkaniu,
- piętro, na którym znajduje się mieszkanie,
- kuchnia w mieszkaniu (widna, ciemna),
- ocena położenia (korzystna, przeciętna),
- standard wykończenia mieszkania (wysoki, przeciętny, niski),
- liczba kondygnacji budynku,
- technologia budowy (tradycyjna, prefabrykowana),
- lokalizacja mieszkania w budynku (dobra, przeciętna).

Determinanty te mają w sposób „poprawny” określać możliwość sprzedaży mieszkania w czasie krótszym niż średni czas sprzedaży nieruchomości mieszkaniowej (Y).

Oceniając jakość otrzymanych modeli dyskryminacyjnych (poprawność formalno-merytoryczną) brano pod uwagę: współczynnik lambda Wilksa, statystykę F , parametr p , współczynnik R kanoniczny. W procesie oceny otrzymanych modeli neuronowych uwzględniono jakość uczenia oraz jakość testowania sieci neuronowych. Zwrócono również szczególną uwagę, czy zbudowane sieci neuronowe nie utraciły zdolności do generalizacji i uogólniania zdobytej wiedzy. Natomiast przy ocenie praktycznej przydatności zbudowanych modeli decyzyjnych została wyznaczona oraz oceniona macierz poprawnych klasyfikacji.

3. Wnioski z przeprowadzonych badań

Omawiając parametry otrzymanych modeli użyto nazw: model 1, model 2, ..., model 6. Nazwy te zostały jednoznacznie przypisane do rodzaju zastosowanych w nich narzędzi obliczeniowych oraz wykorzystanych w procesie ich budowy rodzaju danych (por. rys. 1).

Model 1 z wykorzystaniem analizy dyskryminacyjnej bazujący na danych ilościowych okazał się nieistotny statystycznie. Cechował się parametrem $p < 0,133$.

Model 2 z wykorzystaniem narzędzi sieci neuronowych perceptronowej wielowarstwowej na bazie danych ilościowych uzyskał wartości dopasowania na zbiorze uczącym i testującym na poziomie odpowiednio 0,68 oraz 0,72. Jednakże podczas procesu symulacyjnego jej testowania uzyskiwała ona dla różnych sytuacji opisujących rynek jednakowe wartości dopasowania. Niewątpliwie świadczy to o utracie przez sieć neuronową zdolności do generalizacji i uogólniania zdobytej wiedzy. Utrata zdolności generalizacji i uogólniania zdobytej wiedzy dyskwalifikuje zbudowany model neuronowy do zastosowań praktycznych.

Model 3 dyskryminujący zależności na bazie danych jakościowych uzyskał poprawne wyniki z punktu widzenia jego parametrów (λ Wilksa = 0,77, $F = 8,47$, $p < 0,0006$, kan. $R = 0,48$). Jest to pierwszy model, który po uzyskaniu zadowalających cech jakościowych został poddany weryfikacji w zastosowaniu praktycznym (wyznaczono procent poprawnej klasyfikacji). Macierz klasyfikacji dla zbioru uczącego przedstawiono w tabeli 1, natomiast dla zbioru testowego w tabeli 2.

Tabela 1. Macierz klasyfikacji – model 3 – zbiór danych uczących

Klasa	Procent poprawnie	Grupa 0	Grupa 1
Grupa 0	65,22	15	8
Grupa 1	77,78	8	28
Razem	72,88	23	36

Źródło: opracowanie własne.

Tabela 2. Macierz klasyfikacji – model 3 – zbiór danych testowych

Klasa	Procent poprawnie	Grupa 0	Grupa 1
Grupa 0	0,00	0	8
Grupa 1	100,00	0	31
Razem	79,49	0	39

Źródło: opracowanie własne.

Poddając analizie otrzymane wyniki wnioskujemy, iż model 3 poprawnie działa na zbiorze danych uczących. Natomiast nie potrafi dokonać poprawnej klasyfikacji nieruchomości mieszkaniowych w grupie 0. Wszystkie przypadki są klasyfikowane przez model jako grupa 1.

Model 4 z wykorzystaniem narzędzi sieci neuronowych na bazie danych jakościowych cechował się utratą generalizacji i uogólniania zdobytej wiedzy. Analogicznie jak w modelu 2, wada ta dyskwalifikuje model w zastosowaniach praktycznych.

Model 5 z wykorzystaniem analizy dyskryminacyjnej bazujący na danych mieszanych (ilościowych, jakościowych) okazał się modelem istotnym statystycznie (λ Wilksa = 0,73, $F = 4,80$, $p < 0,0022$, kan. $R = 0,51$). Przeprowadzając weryfikację modelu w zastosowaniach praktycznych, wyznaczono macierz poprawnych klasyfikacji dla zbioru uczącego (tab. 3) oraz zbioru testującego (tab. 4).

Tabela 3. Macierz klasyfikacji – model 5 – zbiór danych uczących

Klasa	Procent poprawnie	Grupa 0	Grupa 1
Grupa 0	69,57	16	7
Grupa 1	83,33	6	30
Razem	77,97	22	37

Źródło: opracowanie własne.

Tabela 4. Macierz klasyfikacji – model 5 – zbiór danych testowych

Klasa	Procent poprawnie	Grupa 0	Grupa 1
Grupa 0	62,50	5	3
Grupa 1	93,55	2	29
Razem	87,18	7	32

Źródło: opracowanie własne.

Model 6 z wykorzystaniem narzędzi sieci neuronowych na bazie danych jakościowych cechował się utratą generalizacji i uogólniania zdobytej wiedzy. Analogicznie jak w modelu 2 oraz 4, wada ta dyskwalifikuje model w zastosowaniach praktycznych.

Podsumowanie

Podsumowując przeprowadzony proces badawczy oraz proces oceny jakości i przydatności zbudowanych modeli decyzyjnych można wnioskować, iż najlepszym modelem w ramach przyjętych kryteriów jest model bazujący na analizie dyskryminacyjnej ze zbiorem danych mieszanych. Zmienne wykorzystane

w modelu to: przeciętny standard wykończenia, technologia tradycyjna, cena za metr kwadratowy, dobra lokalizacja. Odpowiadające im współczynniki standaryzowane, przydatne w interpretacji modelu zostały przedstawione w tabeli 5.

Tabela 5. Współczynniki standaryzowane – analiza dyskryminacyjna – model 5

Zmienne diagnostyczne	Współczynniki standaryzowane
Przeciętny standard wykończenia	-0,889748
Technologia tradycyjna	0,371334
Cena za metr	-0,397808
Lokalizacja dobra	0,326806
Wartość własna	0,355805
Skumulowany procent	1,000000

Źródło: opracowanie własne.

Przeprowadzony proces badawczy pokazał przydatność zastosowania analizy dyskryminacyjnej w procesie parametryzacji rynku nieruchomości mieszkaniowych. Przyjmując za kryterium wyboru praktyczną przydatność zbudowanych modeli analitycznych, lepszym okazał się model na bazie analizy dyskryminacyjnej (w porównaniu z modelem z wykorzystaniem sieci neuronowych). Wobec powyższego model zbudowany z wykorzystaniem narzędzi dyskryminacji może być stosowany jako narzędzie wspomagające podjęcie trafnych decyzji inwestycyjnych zarówno przez deweloperów, jak i indywidualnych sprzedających. Przyjmując za kryterium czas sprzedaży mieszkania (liczbę dni potrzebną na sprzedaż mieszkania), przy określonych parametrach wartościujących mieszkanie można wskazać szanse powodzenia transakcji lub jej brak. Niepokojącym parametrem w zdefiniowanym modelu dyskryminacyjnym jest jego relatywnie słaba zdolność dyskryminacyjna. Wynika ona z trudności pojawiających się w procesie badawczym, tzn. z subiektywizmu kupujących oraz trudności z pozyskaniem danych pierwotnych. W dalszych etapach badawczych, próbując poprawić jakość zbudowanego modelu dyskryminacyjnego, będzie podjęta próba polepszenia jakości danych wejściowych użytych do budowy modelu.

Literatura

- Adamska M., *Diagnoza czynników warunkujących rozwój małych i średnich przedsiębiorstw, w: Od naukowej inspiracji do innowacji w przedsiębiorstwie. Praktyczna aplikacja wiedzy Asystentów innowacji*, red. M. Szewczuk-Stępień, M. Adamska, Instytut Trwałego Rozwoju, Opole 2013.
- Aczel A.D., *Statystyka w zarządzaniu*, WN PWN, Warszawa 2000.

- Mach Ł., *Aproksymacja zjawisk rynkowych narzędziem wspomagającym podejmowania decyzji*, t. 2, red. R. Knosala, Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją, Opole 2009.
- Mach Ł., *Modelowanie logitowe narzędziem wspomagającym podejmowanie decyzji na rynku nieruchomości mieszkaniowych*, w: *Prognozowanie w zarządzaniu firmą*, red. P. Dittmann, „Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu” 2010.
- Osowski S., *Sieci neuronowe do przetwarzania informacji*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2000.
- Panek T., *Statystyczne metody wielowymiarowej analizy porównawczej*, Wyd. Szkoły Głównej Handlowej w Warszawie, Warszawa 2009.
- Stanisz A., *Przystępny kurs statystyki z zastosowaniem Statistica PL na przykładach z medycyny*, t. 3, *Analizy wielowymiarowe*, Statsoft, Kraków 2007.
- Witkowska, D., *Sztuczne sieci neuronowe i metody statystyczne*, Wyd. C.H. Beck, Warszawa 2002.

Application of discrimination analysis and neural networks in the process of assessment of possibilities of selling real estates

Summary. The article presents the possibilities of application of method of discrimination analysis and tools of artificial neuron networks in the process of assessment of possibilities of selling real estates. The author attempts to compare method classified as multidimensional statistical method (discrimination analysis) and the method based on artificial intelligence (neuron networks). The main purpose of this article is to focus on the potential possibilities of a precise time determination needed to sell a given real estate, which is characterized by a defined set of parameters (*i.e.* price, internal area, number of rooms, *etc.*). Appropriately-prepared analytical models based on discrimination analysis and neuron networks are a key determinant influencing improvement of competitiveness of companies, especially in the time of global economic crisis.

Key words: discrimination analysis, neural networks, residential real estate market