

Ewa Putek-Szeląg, Urszula Gieraltowska

Uniwersytet Szczeciński

WYKORZYSTANIE NIEKLASYCZNYCH METOD KLASYFIKACJI DO ANALIZY RYNKU NIERUCHOMOŚCI MIESZKANIOWYCH

Streszczenie: Zasadniczym celem artykułu jest próba wykorzystania nieklasycznych metod klasyfikacji do grupowania lokali mieszkalnych na podstawie wybranych atrybutów, które kształtują wartość nieruchomości. W tym celu wykorzystano drzewa klasyfikacyjne oraz sztuczne sieci neuronowe. Badania empiryczne przeprowadzono, opierając się na transakcjach kupna-sprzedaży z okresu od 1.09.2008 r. do 30.06.2009 r. Otrzymano informacje o 1570 nieruchomościach mieszkaniowych, ale pełną bazę atrybutów uzyskano dla 250 lokali. W artykule zaprezentowano wykorzystanie nieklasycznych metod podziału do klasyfikacji nieruchomości do dwóch oraz czterech grup cenowych.

1. Wstęp

Rynek nieruchomości w Polsce jest jeszcze młodym, nieukształtowanym rynkiem i różni się od rynków rozwiniętych, co jest spowodowane zarówno sytuacją gospodarczą, szczególnym charakterem tego rynku, jak i obowiązującym stanem prawnym [Kucharska-Stasiak 2006]. Przystąpienie do UE zbiegło się ze wzrostem popytu na rynku nieruchomości, co przełożyło się na wzrost cen gruntów, domów, mieszkań i materiałów budowlanych. Zjawisko to spowodowało zainteresowanie inwestorów nieruchomościami jako przedmiotem, w który można dokonać lokaty posiadanych środków.

Problematyka rynku nieruchomości stanowi interesujące pole badawcze dla analityków pragnących zbudować modele rynkowe, które w sposób efektywny pozwolą oszacować wartość nieruchomości na podstawie znanych atrybutów. Trudności powstałe w trakcie procesu analitycznego w odniesieniu do rynku nieruchomości wynikają głównie z: utrudnionego dostępu do informacji, rzadkości transakcji, nieracjonalnego zachowania kupujących i sprzedających, dużej różnicy cen pomiędzy podobnymi, ale mało porównywalnymi nieruchomościami, różnorodności nieruchomości, sposobu użytkowania.

Zasadniczym celem artykułu jest próba wykorzystania nieklasycznych metod klasyfikacji do grupowania lokali mieszkalnych na podstawie wybranych atrybutów, które kształtują wartość nieruchomości. Spośród wielu metod zaproponowano wykorzystanie drzew klasyfikacyjnych oraz sztucznych sieci neuronowych.

2. Opis grupy badawczej

Badania empiryczne przeprowadzono, opierając się na danych rzeczywistych pochodzących ze szczecińskiego rynku nieruchomości. Dane dotyczą transakcji z okresu od 1.09.08 r. do 30.06.09 r. Otrzymano informacje o 1570 nieruchomościach mieszkaniowych, ale pełną bazę atrybutów otrzymano dla 250 lokali.

Wszystkie mieszkania opisane zostały przez następujące zmienne: x_1 – cena transakcyjna (w zł), x_2 – powierzchnia mieszkania (w m^2), x_3 – cena 1 m^2 (w zł), x_4 – obręb (0 – dzielnica mało atrakcyjna, 1 – umiarkowanie atrakcyjna, 2 – modna), x_5 – rodzaj nabywanego prawa (0 – spółdzielcze własnościowe prawo, 1 – użytkowanie wieczyste, 2 – własnościowe), x_6 – liczba pokoi, x_7 – rok budowy (0 – zabudowa przedwojenna, 1 – lata 60.-80., 2 – lata 90., 3 – zabudowa współczesna (po roku 2000)), x_8 – charakterystyka zabudowy (0 – zabudowa willowa, 1 – zabudowa niska, 2 – zabudowa wysoka), x_9 – kondygnacja – numer piętra, na którym znajduje się mieszkanie.

Ze względu na fakt, że cena transakcyjna wyrażająca wartość nieruchomości mieszkaniowej w znacznej mierze uzależniona jest od jej powierzchni, zatem klasyfikacja lokali na jej podstawie będzie zdominowana przez ten właśnie atrybut. W dalszych analizach przyjęto, że obiektywnym czynnikiem pozwalającym na klasyfikację obiektów będzie cena 1 m^2 . Analizowane nieruchomości są istotnie zróżnicowane pod względem omawianych atrybutów. W tabelach 1-3 przedstawiono kształtowanie się przeciętnych cen 1 m^2 powierzchni w zależności od położenia nieruchomości (przedstawiono te dzielnice, w których dokonywano dużej liczby transakcji), liczby pokoi oraz wieku budynku, w którym położona jest nieruchomość.

Tabela 1. Kształtowanie się cen 1 m^2 nieruchomości mieszkalnych w wybranych dzielnicach Szczecina w okresie 1.09.08-30.06.09

Dzielnica – obręb	Średnia cena 1 m^2	Odchylenie standardowe
Warszewo	5168,38	1571,09
Stare Miasto	4977,59	1192,07
Krzekowo	4952,65	1613,22
Gumieńce	4928,73	1035,71
Pogodno	4613,29	1312,68
Pomorzany	4297,27	656,87
Śródmieście	4519,76	818,06
Niebuszewo	4504,27	785,71
Słoneczne	4276,26	1228,50
Dąbie	3969,81	1021,00
Załom	3529,31	603,34
Podjuchy	3505,50	1346,50

Źródło: obliczenia własne.

Tabela 2. Kształtowanie się cen 1 m² nieruchomości mieszkalnych w Szczecinie według liczby pokoi w okresie 1.09.08-30.06.09

Liczba pokoi	1	2	3	4	5
Średnia cena 1 m ²	4821,96	4390,16	4405,70	3669,11	3162,63
Odchylenie standardowe	1148,62	1092,52	1015,71	2039,98	1522,42

Źródło: obliczenia własne.

W dalszych rozważaniach przyjęto, że głównym kryterium podziału będzie mediana wartości (cena 1 m² charakteryzuje się silnym asymetrycznym rozkładem), co pozwoliło na wyodrębnienie w miarę jednorodnych liczebnie grup. Po wstępnej analizie cen 1 m² wszystkich mieszkań przyjęto, że rozważania będą przebiegały dwuetapowo. W pierwszym etapie przeprowadzona zostanie analiza jedynie dla dwóch grup nieruchomości, a według kryterium wyodrębnienia podziału dychotomicznego struktura mieszkań miała postać: grupa I – cena 1 m² poniżej 4530,74 zł, grupa II – powyżej 4530,74 zł.

Tabela 3. Kształtowanie się cen 1 m² nieruchomości mieszkalnych w Szczecinie według wieku budynku w okresie 1.09.08-30.06.09

Wiek budynku	Przedwojenny	Lata 60.-80.	Lata 90.	Lata 2000-2009
Średnia cena 1 m ²	4117,21	4510,52	4958,86	5141,82
Odchylenie standardowe	1236,12	1052,61	941,59	1290,82

Źródło: obliczenia własne.

W drugim etapie do dalszych analiz zostały wyodrębnione 4 grupy mieszkań, przy czym grupy cenowe zostały przyjęte na poziomie kwartyła pierwszego (Q_{1,4}), mediany (M) oraz kwartyła trzeciego (Q_{3,4}). Wyodrębniono grupy cenowe: I (cena 1 m² poniżej 3993,34 zł), II (cena pomiędzy (3993,34; 4530,74) zł), III (cena pomiędzy (4530,74; 5113,64) zł) i IV (cena powyżej 5113,64 zł).

3. Metodologia przeprowadzonych badań

Niewątpliwą wadą wielu klasycznych metod klasyfikacji obiektów jest konieczność przyjęcia założenia dotyczącego wielowymiarowego rozkładu normalnego zmiennych diagnostycznych. W praktyce spełnienie tego założenia staje się dużym ograniczeniem i może w znacznym stopniu rzutować na jakość uzyskanych klasyfikacji. Rozwiązaniem problemu może być wykorzystanie metod, których stosowanie nie jest obwarowane tak surowymi założeniami. Do takich metod niewątpliwie można zaliczyć metody nieparametryczne oparte na drzewach klasyfikacyjnych czy sieciach neuronowych, które pozwalają na budowę modelu bez znajomości postaci rozkładów cech oraz związków między nimi.

Drzewa klasyfikacyjne przedstawiają proces podziału zbioru obiektów na jednorodny klasy na podstawie wartości cech obiektów (liście odpowiadają klasom, a krawędzie drzewa reprezentują wartości cech, na podstawie których dokonano podziału). Węzły reprezentują podzbiory, które mogą ulegać dalszym podziałom (w poszczególnych węzłach określone są reguły klasyfikacyjne). Drzewa klasyfikacyjne są wykorzystywane w naukach stosowanych, ponieważ za pomocą czytelnego i prostego obrazu graficznego ułatwiają interpretację skomplikowanych zjawisk wielowymiarowych¹.

Zasadniczym celem jest utworzenie drzewa o jak najmniejszej liczbie węzłów, aby otrzymać proste reguły klasyfikacyjne. Charakterystyczna dla drzew klasyfikacyjnych jest ich hierarchiczna natura oraz elastyczność. Wśród algorytmów obliczeniowych najczęściej stosowanymi są: C&RT, QUEST oraz CRUISE. Do tworzenia drzew klasyfikacyjnych wykorzystać można m.in. moduł drzewa klasyfikacyjne z pakietu STATISTICA 8.0.

Sieci neuronowe² stosowane są do rozwiązywania wielu problemów badawczych związanych z klasyfikacją oraz predykcją w różnorodnych dziedzinach. Zaletami sieci neuronowych w porównaniu z metodami klasycznymi są przede wszystkim nieliniowy charakter oraz brak założeń dotyczący danych wejściowych. Sztuczne sieci neuronowe w sposób szybki i efektywny przetwarzają olbrzymie informacje danych, często chaotycznych, niekompletnych, a nawet sprzecznych, funkcjonując nawet w momencie uszkodzenia lub zerwania połączenia.

Badania przeprowadzono z wykorzystaniem pakietu STATISTICA 8.0 (moduł automatyczne sieci neuronowe). W badaniach wykorzystano architekturę sieciową opartą na koncepcji perceptronu wielowarstwowego (MLP – *Multilayer Perceptron*).

4. Mierniki dokładności klasyfikacji

Ocena jakości klasyfikacji dokonywana jest przez porównanie „zawartości” poszczególnych klas ze stanem rzeczywistym. Podstawową miarą jest ogólny błąd klasyfikacji, który pozwala określić, jaka część obiektów została niepoprawnie rozpoznana na skutek zastosowania wybranej reguły klasyfikacyjnej. Błąd ogólny ma postać³:

$$E = \frac{Z}{n} \cdot 100\%, \quad (1)$$

¹ Szerzej o własnościach sieci neuronowych w pracach m.in. E. Gatnara [1999; 2001].

² Problematyka sztucznych sieci neuronowych jest szeroko omawiana w literaturze, ale nie stanowi to istoty artykułu. Wystarczy wspomnieć o pracach: W.S McCullocha, W. Pittsa, C.M. Bishopa, R. Tadeusiewicza, J. Korbicza i innych. Krótki przegląd prac, jak również przykładowe zastosowanie sieci neuronowych w badaniach ekonomicznych znaleźć można m.in. w pracy D. Witkowskiej [2002] czy M. Lasek [2002].

³ Zob. [Witkowska 2002, s. 92-93].

gdzie: n – liczba obiektów objętych badaniem, Z – liczba błędnie zaklasyfikowanych obiektów (pochodzących z klasy C_r , a zaklasyfikowanych do C_s i odwrotnie, przy $r \neq s$).

Można również analizować rozkład błędnie rozpoznanych obiektów w poszczególnych klasach. Częstkowy błąd klasyfikacji ma postać:

$$E_k = \frac{Z_k}{n_k} \cdot 100\%, \quad (2)$$

gdzie: Z_k – liczba błędnie zaklasyfikowanych obiektów pochodzących z klasy C_k ,
 n_k – liczebność k -tej klasy.

W przypadku grupowania dla dwóch klas można wyznaczyć tzw. błędy pierwszego i drugiego rodzaju. W przypadku, kiedy klasyfikacja nie jest dychotomiczna, możliwe jest również wyznaczenie błędu klasyfikacji do klasy innej niż sąsiednia.

5. Wyniki klasyfikacji przy wykorzystaniu drzew klasyfikacyjnych

W procesie budowy drzew klasyfikacyjnych zastosowano algorytm C&RT⁴. Przyjęto proporcjonalne do liczebności prawdopodobieństwo *a priori* oraz równe koszty błędnej klasyfikacji. Kryterium była reguła przycinania przy błędzie złej klasy-

Tabela 4. Charakterystyki kolejnych drzew otrzymanych algorytmem C&RT

Nr	Liczba węzłów końcowych	Koszty sprawdzianu krzyżowego	Błąd standardowy kosztu	Koszty resubstytucji	Złożoność drzewa
1	31	0,440000	0,035100	0,220000	0,000000
2	27	0,440000	0,035100	0,225000	0,001250
3	23	0,440000	0,035100	0,235000	0,002500
4	16	0,425000	0,034955	0,255000	0,002857
*5	11	0,430000	0,035007	0,280000	0,005000
6	9	0,490000	0,035348	0,295000	0,007500
7	7	0,490000	0,035348	0,315000	0,010000
8	5	0,490000	0,035348	0,340000	0,012500
9	4	0,525000	0,035311	0,355000	0,015000
10	2	0,510000	0,035348	0,410000	0,027500
11	1	0,480000	0,035327	0,480000	0,070000

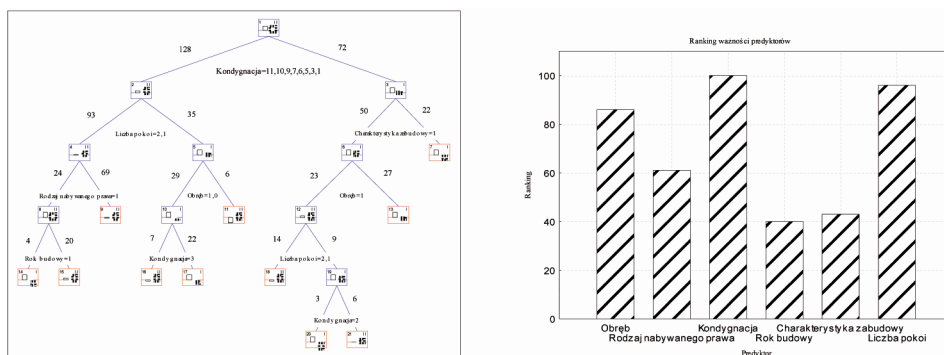
Źródło: opracowanie własne na podstawie STATISTICA 8.0.

⁴ Testowano również inne procedury drzew klasyfikacyjnych dostępnych w programie Statistica. Niestety dokładność przeprowadzanych klasyfikacji była dużo niższa lub porównywalna z prezentowanymi wynikami, lecz struktura drzew była bardziej rozbudowana, dlatego wyniki nie zostały zaprezentowane w pracy.

fikacji. Z analizowanej bazy losowo wyodrębniono próbę uczącą (80%) i testową (20%).

W pierwszej kolejności dokonano klasyfikacji nieruchomości na podstawie dwóch grup cenowych nieruchomości. W tabeli 4 przedstawiono statystyki drzew wyznaczonych za pomocą algorytmu C&RT.

Najlepszym drzewem było drzewo nr 5 złożone z jedenastu węzłów końcowych. Graficzna prezentacja drzewa została przedstawiona na rys. 1 (część lewa). Należy zauważyć, że wśród czynników, które miały największy wpływ na kształtowanie się cen 1 m² nieruchomości mieszkalnej, były: kondygnacja, liczba pokoi oraz dzielnica. Pozostałe atrybuty również miały wpływ na cenę, jednak już mniejszy. Ranking atrybutów znajduje się na rys. 1 (część prawa).



Rys. 1. Graficzna prezentacja drzewa klasyfikacyjnego oraz ranking ważności predyktorów przy podziale nieruchomości na dwie klasy

Źródło: opracowanie własne na podstawie STATISTICA 8.0.

W tabeli 5 przedstawiono wyniki klasyfikacji. W zbiorze uczącym dokładność klasyfikacji wynosiła 72%. Nieco gorzej było w odniesieniu do zbioru testowego, ponieważ dokładność wynosiła jedynie 60%. W obu przypadkach zbudowane drzewo dużo lepiej klasyfikowało obiekty w drugiej grupie, w której przeciętna cena 1 m² jest powyżej wartości mediany. Może to sugerować, że duża część cen

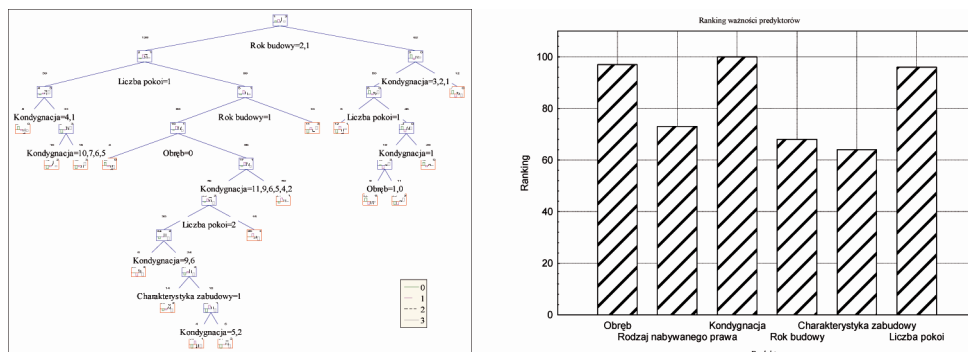
Tabela 5. Ocena jakości klasyfikacji nieruchomości dla grupy uczącej i testowej przy podziale na dwie grupy cenowe

Grupa cenowa	Zbiór uczący		Zbiór testowy	
	I	II	I	II
I	59	19	12	6
II	37	85	14	18
E _k [%]	38,54	18,27	53,85	25
E [%]	28		40	

Źródło: opracowanie własne na podstawie STATISTICA 8.0.

transakcyjnych w badanym okresie jest zaniżona w stosunku do ceny, jaka powinna zostać osiągnięta. Wydaje się to powiązane ze źródłem pochodzenia danych. W wielu przypadkach strony dokonujące transakcji kupna-sprzedaży nieruchomości celowo zaniżają wartość nieruchomości, aby zapłacić niższy podatek.

W przypadku klasyfikacji nieruchomości na podstawie czterech grup cenowych najlepszym drzewem było drzewo składające się z 16 węzłów końcowych⁵. Graficzna prezentacja drzewa została zaprezentowana na rys. 2.



Rys. 2. Graficzna prezentacja drzewa klasyfikacyjnego oraz ranking ważności predyktorów przy podziale nieruchomości na cztery klasy

Źródło: opracowanie własne na podstawie STATISTICA 8.0.

Należy zwrócić uwagę, że struktura drzewa nie jest już tak czytelna. Spośród zaproponowanych atrybutów najistotniejsze są ponownie: kondygnacja, dzielnica oraz liczba pokoi, przy czym pozostałe charakterystyki były również ważne. W tabeli 6 przedstawiono wyniki klasyfikacji nieruchomości przy podziale na cztery grupy cenowe. W przypadku podziału nieruchomości mieszkaniowych na 4 klasy uzyskano dużo gorsze wyniki klasyfikacji. Dla grupy uczącej otrzymano dokładność na poziomie 56%, a w przypadku grupy testowej – jedynie 32%.

Należy zauważyć, że nie wszystkie błędnie zakwalifikowane obiekty znalazły się w sąsiednich grupach. W przypadku zbioru uczącego odsetek błędnie zaklasyfikowanych obiektów do klasy innej niż sąsiednia wyniósł 21%, a w przypadku grupy testowej – 20%.

Na podstawie otrzymanych wyników można stwierdzić, że drzewa klasyfikacyjne dają dość przeciętne wyniki klasyfikacji obiektów na podstawie ceny 1 m². Należy jednak podkreślić, że szczególnie w ostatnim okresie, kiedy na świecie mamy do czynienia z kryzysem gospodarczym, olbrzymią rolę w kształtowaniu cen (jak również popytu i podaży) odgrywały czynniki psychologiczne i ekono-

⁵ W przypadku innych algorytmów drzew klasyfikacyjnych otrzymane wyniki były dużo gorsze, ponieważ obiekty były klasyfikowane tylko do dwóch grup cenowych.

Tabela 6. Ocena jakości klasyfikacji nieruchomości dla grupy uczącej i testowej przy podziale na cztery grupy cenowe

Grupa cenowa	Zbiór uczący				Zbiór testowy			
	I	II	III	IV	I	II	III	IV
I	31	6	13	8	5	3	3	1
II	8	32	14	10	5	6	7	2
III	4	8	27	3	1	2	5	2
IV	0	7	7	22	2	1	5	0
E_k [%]	27,9	39,6	55,7	48,8	61,5	50	75	100
E [%]	44				68			

Źródło: opracowanie własne na podstawie STATISTICA 8.0.

miczno-społeczne. Zmodyfikowane procedury bankowe miały duży wpływ na popyt na nieruchomości, przez co ceny 1 m^2 nieruchomości znacznie odbiegały od cen sprzed kryzysu.

6. Wyniki klasyfikacji przy wykorzystaniu sieci neuronowych

Do klasyfikacji nieruchomości na podstawie sztucznych sieci neuronowych wykorzystano sieci MLP. W przypadku klasyfikacji na podstawie dwóch klas najlepsza otrzymana sieć miała strukturę $6 - 10 - 2$, co oznacza, że sieć składała się z dziesięciu neuronów ukrytych (budowane sieci różniły się przede wszystkim funkcjami aktywacji – w przypadku najlepszej sieci była to funkcja hiperboliczna dla ukrytych i logistyczna dla wyjścia). Wyniki klasyfikacji przedstawiono w tab. 7.

Tabela 7. Ocena jakości klasyfikacji nieruchomości dla grupy uczącej i testowej przy podziale na dwie grupy cenowe na podstawie sieci neuronowych

Grupa cenowa	Zbiór uczący		Zbiór testowy	
	I	II	I	II
I	60	25	16	9
II	36	79	10	15
E_k [%]	37,5	24,04	38,46	37,5
E [%]	30,5		38	

Źródło: opracowanie własne na podstawie STATISTICA 8.0.

Dokładność klasyfikacji na podstawie zbioru uczącego wynosiła 69,5%, natomiast na podstawie zbioru testowego była nieznacznie niższa i wynosiła 62%. Podobnie jak w przypadku drzew klasyfikacyjnych wyższą dokładność otrzymano dla obiektów w drugiej grupie, których przeciętna cena 1 m^2 była powyżej wartości mediany. Podczas badania wrażliwości poszczególnych atrybutów okazało się, że wszystkie wyróżnione czynniki miały wpływ na jakość sieci, przy czym najbardziej istotne to dzielnica, liczba pokoi i rok budowy.

W przypadku podziału na cztery grupy cenowe otrzymano sieć o strukturze 6 – 7 – 4 (funkcje aktywacji to funkcja hiperboliczna (dla neuronów ukrytych) oraz softmax dla wyjścia). Dokładność klasyfikacji przedstawiono w tab. 8.

Tabela 8. Ocena jakości klasyfikacji nieruchomości dla grupy uczącej i testowej przy podziale na cztery grupy cenowe przy wykorzystaniu sieci neuronowych

Grupa cenowa	Zbiór uczący				Zbiór testowy			
	I	II	III	IV	I	II	III	IV
I	24	2	7	5	4	1	0	0
II	5	25	11	7	1	7	5	0
III	8	20	32	9	5	4	12	2
IV	6	6	11	22	3	1	3	2
E_k [%]	44,19	52,83	47,54	48,84	69,23	46,15	40	50
E [%]	48,5				50			

Źródło: opracowanie własne na podstawie STATISTICA 8.0.

W przypadku podziału nieruchomości mieszkaniowych na 4 klasy uzyskano wyniki klasyfikacji zbliżone do wyników uzyskanych przy wykorzystaniu drzew. Dla grupy uczącej otrzymano dokładność na poziomie 51,5%, a w przypadku grupy testowej – 50%. Błędy klasyfikacji do grupy innej niż sąsiednia wynosiły odpowiednio 19,5 oraz 18%. Podobnie jak przy podziale na dwie klasy analiza wrażliwości wskazała, że najistotniejszy wpływ na cenę 1 m² nieruchomości miały obręb, liczba pokoi oraz rok budowy nieruchomości.

7. Podsumowanie i wnioski

Z przeprowadzonych badań wynika, że lepsze wyniki klasyfikacji mierzone dokładnością podziału uzyskano dla mniejszej liczby klas (niezależnie od przyjętej nieklasycznej metody klasyfikacji), przy czym nieznacznie wyższą dokładność uzyskano przy wykorzystaniu drzew klasyfikacyjnych. Wpływ poszczególnych atrybutów na wartość nieruchomości ulegał zmianie w zależności od liczby klas wstępnego podziału. Jednak uzyskane wyniki mogą świadczyć o tym, że występują jeszcze inne czynniki, które silnie wpływają na cenę 1 m², ale nie zostały one uwzględnione w analizie. W przypadku podziału nieruchomości na cztery klasy błąd klasyfikacji do grupy innej niż sąsiednia jest dość wysoki (ok. 20%), co może oznaczać, że zdarzają się obiekty w sąsiednich grupach, które nie wykazują do siebie podobieństwa. Na rynku nieruchomości często się zdarza, że na cenę 1 m² wpływ mają również czynniki psychologiczne, które trudne są do analizy.

Zasadniczą przyczyną tak niskiej dokładności prezentowanych klasyfikacji jest bardzo niska jakość danych wejściowych. Wyższa jakość danych oraz poprawa ich wiarygodności powinna wpłynąć dodatnio na dokładności klasyfikacji.

Uzyskane wyniki mogą stanowić cenne źródło informacji dla instytucji państwowych, m.in. urzędów skarbowych. Przedstawione procedury mogą stanowić podstawę weryfikacji cen transakcyjnych zawieranych na rynkach lokalnych i sprawdzenia, czy nie zostały one zaniżone w celu uniknięcia wyższych kosztów transakcyjnych.

Literatura

- Gatnar E., *Nieparametryczna metoda dyskryminacji i regresji*, Wydawnictwo Naukowe PWN, Warszawa 2001.
- Gatnar E., *Drzewa klasyfikacyjne: nieparametryczna metoda dyskryminacji obiektów*, „Badania Operacyjne i Decyzyjne” 1999 nr 1.
- Kucharska-Stasiak E., *Nieruchomość w gospodarce rynkowej*, Wydawnictwo Naukowe PWN, Warszawa 2006.
- Lasek M., *Data Mining. Zastosowania w analizach i ocenach klientów bankowych*, Biblioteka Menadżera i Bankowca, Warszawa 2002.
- Witkowska D., *Sztuczne sieci neuronowe i metody statystyczne. Wybrane zagadnienia finansowe. Studia ekonomiczne*, Wydawnictwo C.H. Beck, Warszawa 2002.

THE USE OF NON-CLASSICAL CLASSIFICATION METHODS FOR ANALYZING THE REAL ESTATE MARKET

Summary: The primary objective of the article is an attempt to use non-classical classification methods to group dwellings on the basis of selected attributes that shape the value of the property. For this purpose classification trees and artificial neural networks have been used.

Empirical studies have been conducted on the basis of sale/purchases transactions during the period from 01.09.08 to 30.06.09. Information has been received for 1570 residential real estate, but the full database of attributes has been obtained for 250 apartments. In the paper, we present the results of using the non-classical methods to classify real estate to two or four price groups.