

Polaryzacja konsumpcji w gospodarstwach domowych

We współczesnym świecie konsumpcja dóbr i usług podlega intensywnym przemianom. Z jednej strony obserwuje się pauperyzację znaczącej części gospodarstw domowych, z drugiej zaś uwagę wielu badaczy przyciąga problem nadmiernej konsumpcji. Konsumpcjonizm obserwowany jest nie tylko w krajach wysoko rozwiniętych, ale także w krajach o niższym poziomie rozwoju.

MATERIAŁ I METODYKA BADANIA

Celem badania było wyodrębnienie względnie jednorodnych grup gospodarstw domowych należących do określonej grupy, ze względu na poziom dochodów do dyspozycji na osobę. Do realizacji celu zastosowano metodę drzew klasyfikacyjnych i regresyjnych CART.

Rozpatrywany problem badawczy został opracowany na podstawie indywidualnych danych z badania budżetów gospodarstw domowych GUS za rok 2006, którym objęto 37508 gospodarstw domowych.

Do przeprowadzenia doświadczenia wykorzystano drzewa klasyfikacyjne i regresyjne, które służą do konstrukcji modeli predykcyjnych i deskryptywnych. Drzewo jest graficznym modelem powstałym w wyniku rekurencyjnego podziału zbioru obserwacji X na n rozłącznych podzbiorów $X_1, X_2, X_3, \dots, X_n$. Celem budowy modelu jest uzyskanie podzbiorów maksymalnie homogenicznych względem zmiennej zależnej. Jest to proces wieloetapowy, który w każdej kolejnej fazie może wykorzystywać inną zmienną niezależną. Na każdym etapie analizowane są wszystkie predyktory i wybierany jest tylko ten, który zapewnia najlepszy podział węzła, tzn. wydziela najbardziej jednorodne podzbiory (Łapczyński, 2003).

Pierwszym etapem w analizie danych za pomocą metod drzewkowych jest wybór reguły podziału. Do przeprowadzenia badania wybrano metodę CART oraz regułę podziału — indeks Giniego. Metoda CART jest najbardziej zaawansowaną metodą konstruowania drzew klasyfikacyjnych. Jej początek datuje się na lata 80. ub. stulecia. Swoją popularność zdobyła dzięki kilku cechom: obsługuje zbiory obserwacji o złożonej, wielowymiarowej strukturze, jest niewrażliwa na występowanie obserwacji odstających, wykorzystuje każdą kombinację zmiennych ciągłych i kategoryalnych. Może ona być efektywnie wykorzystywana w zbiorach danych charakteryzujących się licznymi brakami zmiennych niezależnych oraz samych zmiennych w różnych częściach drzewa, odkrywając kontekst interakcji i zależności między zmiennymi (Łapczyński, Sagan, 2008).

Kolejne etapy w budowie drzew klasyfikacyjnych to ustalenie prawdopodobieństwa *a priori* występowania klas oraz wybór kryterium stopu. W przeprowadzonym doświadczeniu wybrano prawdopodobieństwo *a priori* szacowane, a minimalną licznosc w węźle końcowym w pierwszym doświadczeniu ustalono na 200 przypadków, a w drugim — na 50 przypadków.

Początek każdego drzewa stanowi cały zbiór obserwacji, który jest następnie dzielony na dwa (binarne drzewa) lub więcej podzbiorów (dowolne drzewa). Dzielony zbiór nazywamy węzłem macierzystym, zaś wydzielone podzbiory — węzłami potomkami. W kolejnym etapie węzeł potomek ulega kolejnemu podziałowi, przez co staje się węzłem macierzystym dla drugiego etapu. Węzeł, który pozostaje bez zmian, staje się węzłem końcowym, zwanym także liściem. Liczba liści określa wielkość drzewa, natomiast głębokość drzewa to liczba krawędzi między wierzchołkiem a najbardziej odległym liściem.

W pierwszej analizie zmienną zależną była zmienna zero-jedynkowa, gdzie „0” oznaczało gospodarstwa domowe z dochodami do dyspozycji poniżej 650 zł na osobę, a „1” — gospodarstwa domowe z dochodami do dyspozycji powyżej 650 zł na osobę. Dochód do dyspozycji według definicji GUS, to dochód rozporządzalny pomniejszony o pozostałe wydatki, przeznaczony jest na wydatki — towary i usługi konsumpcyjne oraz przyrost oszczędności. Wartość 650 zł została ustalona na podstawie wysokości minimum socjalnego w 2006 r. dla gospodarstw pracowniczych (tabl. 1).

TABL. 1. WYSOKOŚĆ MINIMUM SOCJALNEGO W 2006 R. W GOSPODARSTWACH PRACOWNICZYCH WEDŁUG LICZBY OSÓB ORAZ SKŁADU OSOBOWEGO W ZŁ

Wyszczególnienie	Gospodarstwa domowe					
	1-osobowe (M lub K)	2 (M+K)	3 (M+K+Dml)	3 (K+M+Dst)	4 (K+M+Dml+Dst)	5-osobowe (K+M+Dml+2xDst)
Wysokość minimum socjalnego	802	1318	1958	2027	2506	3056

U w a g a. M — mężczyzna w wieku od 25 do 60 lat; K — kobieta w wieku od 25 do 60 lat; Dml — dziecko młodsze w wieku od 4 do 6 lat; Dst — dziecko starsze w wieku od 13 do 15 lat.

Ź r ó d ł o: strony internetowe IPiSS.

Minimum socjalne to kategoria, która określa koszty utrzymania gospodarstw domowych poprzez uwzględnienie podstawowych potrzeb bytowo-konsumpcyjnych. Wzorzec minimum socjalnego to model zaspokajania potrzeb na ogólnie niskim poziomie, aczkolwiek jeszcze wystarczającym do reprodukcji sił witalnych na każdej fazie biologicznego rozwoju człowieka, do posiadania i wychowania dzieci i do utrzymania więzi ze społeczeństwem (Kurowski, 2002).

W drugim badaniu zmienną zależną była zmienna zero-jedynkowa, gdzie „0” oznaczało gospodarstwa domowe, w których średni dochód do dyspozycji był równy bądź przekraczał 2 tys. zł na osobę, a „1” oznaczała gospodarstwa domowe, w których średni dochód do dyspozycji wynosił poniżej 2 tys. zł na osobę.

Zmiennymi niezależnymi (predyktorami) w obu analizach były:

- klasa miejscowości zamieszkania — gospodarstwa domowe zlokalizowane w miastach (z liczbą osób: ponad 500 tys., od 200 tys. do 500 tys., od 100 tys. do 200 tys., od 20 tys. do 100 tys., poniżej 20 tys.) oraz gospodarstwa domowe zlokalizowane na wsiach,
- liczba osób w gospodarstwie domowym,
- wiek głowy gospodarstwa domowego,
- wykształcenie głowy gospodarstwa domowego (podstawowe, zasadnicze zawodowe, średnie zawodowe, średnie ogólnokształcące, wyższe),
- liczba dzieci na utrzymaniu do 26 roku życia w gospodarstwie domowym.

Ze zbioru danych wyselekcjonowano gospodarstwa domowe pracowników i na tej grupie przeprowadzono doświadczenia.

Analiza statystyczna wykonana została dzięki wykorzystaniu programu statystycznego *Statistica 8,0*.

WYNIKI BADAŃ WŁASNYCH

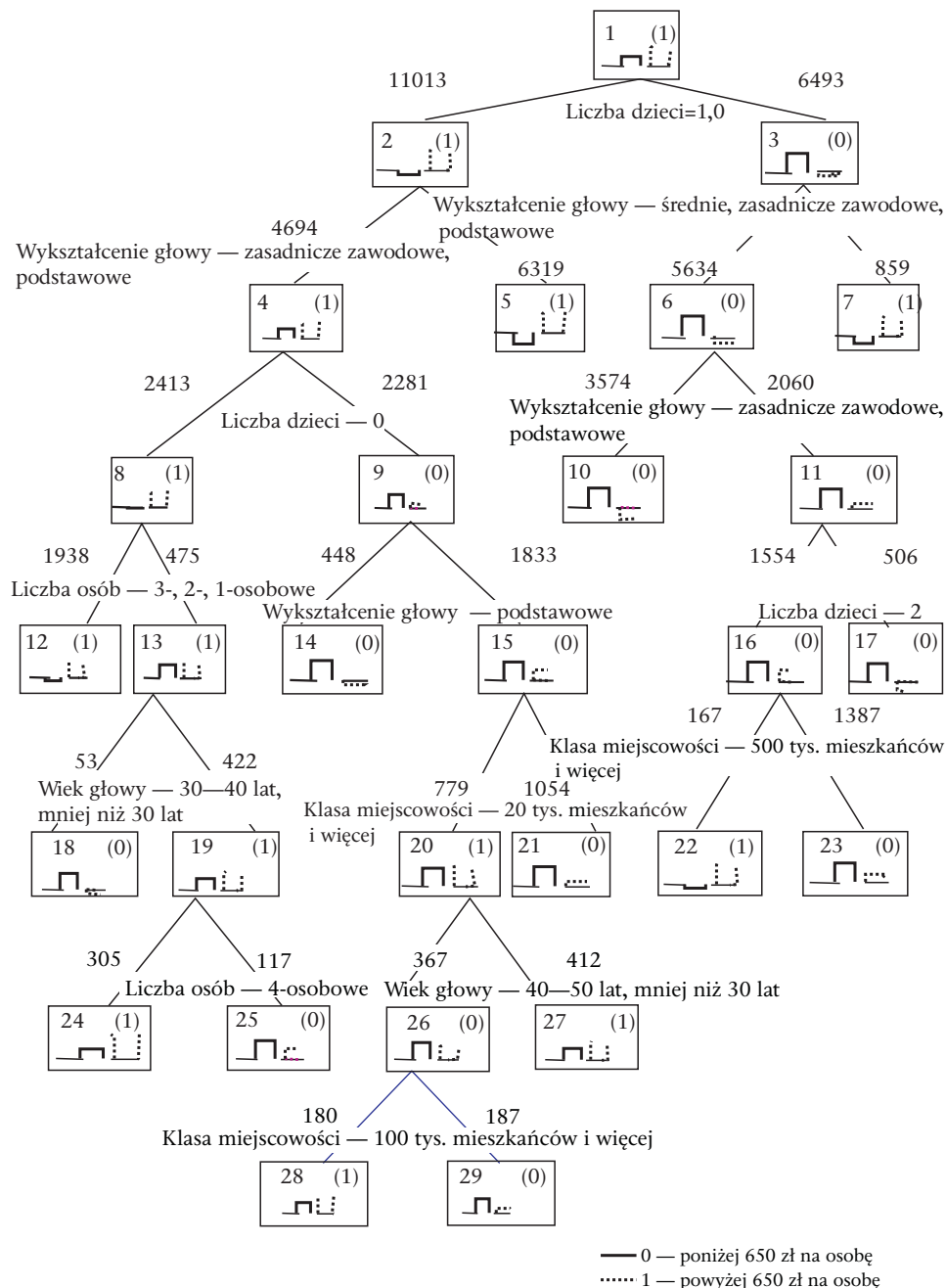
W analizowanym zbiorze GUS gospodarstwa domowe pracowników stanowiły 46,7% badanych ogółem. W ok. 43% gospodarstw domowych dochody do dyspozycji nie przekraczały 650 zł na osobę.

W pierwszym modelu wielkość drzewa, czyli liczba liści wyniosła 15, a głębokość — 7 (schemat 1). Zatem otrzymano 15 węzłów końcowych, czyli 15 zdań warunkowych, za pomocą których można opisać analizowane gospodarstwa domowe.

Dane przy kolejnych węzłach końcowych można zinterpretować następująco:

- V węzeł końcowy — jeśli w gospodarstwie domowym nie było dzieci bądź było jedno dziecko na utrzymaniu do 26 roku życia, a głowa gospodarstwa domowego reprezentowała wykształcenie wyższe niż podstawowe, to gospodarstwo takie pozostawało w grupie gospodarstw z dochodami powyżej 650 zł na osobę, z prawdopodobieństwem 84%;
- VII węzeł końcowy — jeśli w gospodarstwie domowym było przynajmniej dwoje dzieci na utrzymaniu, a głowa gospodarstwa domowego reprezentowała wyższe wykształcenie, to gospodarstwo to pozostawało w grupie gospodarstw z dochodami powyżej 650 zł na osobę, z prawdopodobieństwem 77%;
- X węzeł końcowy — jeśli w gospodarstwie domowym było przynajmniej dwoje dzieci na utrzymaniu, a głowa gospodarstwa domowego reprezentowała podstawowy lub zasadniczy zawodowy poziom wykształcenia, to gospodarstwo takie pozostawało w grupie gospodarstw z dochodami poniżej 650 zł na osobę, z prawdopodobieństwem 83%;
- XII węzeł końcowy — jeśli głowa gospodarstwa domowego miała wykształcenie podstawowe lub zasadnicze zawodowe, w gospodarstwie domowym nie było dzieci na utrzymaniu i składało się ono z jednej, dwu lub trzech osób, to gospodarstwo to pozostawało w grupie gospodarstw z dochodami powyżej 650 zł na osobę, z prawdopodobieństwem 72%;

Schemat 1. DRZEWO KLASYFIKACYJNE CART GOSPODARSTW DOMOWYCH PRACOWNIKÓW Z DOCHODAMI DO DYSPONUCJI NA OSOBĘ PONIŻEJ (0) LUB POWYŻEJ (1) 650 ZŁ NA OSOBĘ MIESIĘCZNIE



Ź r ó ł o: obliczenia własne na podstawie danych z badania budżetów gospodarstw domowych GUS.

- XIV węzeł końcowy — jeśli w gospodarstwie domowym było jedno dziecko na utrzymaniu, a głowa gospodarstwa domowego reprezentowała podstawowy poziom wykształcenia, to gospodarstwo takie pozostawało w grupie gospodarstw z dochodami dyspozycyjnymi poniżej 650 zł na osobę, z prawdopodobieństwem 73%;
- XVII węzeł końcowy — jeśli w gospodarstwie domowym było więcej niż dwoje dzieci na utrzymaniu, a głowa gospodarstwa domowego reprezentowała wykształcenie średnie ogólnokształcące lub zawodowe, wówczas gospodarstwo to pozostawało w grupie gospodarstw z dochodami poniżej 650 zł, z prawdopodobieństwem 79%;
- XVIII węzeł końcowy — jeśli w gospodarstwie domowym nie było dzieci na utrzymaniu, gospodarstwo domowe było przynajmniej czteroosobowe, a wiek głowy gospodarstwa domowego nie przekraczał 40 roku życia, to gospodarstwo to pozostawało w grupie gospodarstw z dochodami dyspozycyjnymi poniżej 650 zł na osobę, z prawdopodobieństwem 72%;
- XXI węzeł końcowy — jeśli w gospodarstwie domowym było jedno dziecko na utrzymaniu, głowa gospodarstwa miała poziom wykształcenia zasadniczy zawodowy, a gospodarstwo domowe zlokalizowane było na wsi albo w mieście z liczbą mieszkańców poniżej 20 tys., to gospodarstwo takie pozostawało w grupie gospodarstw z dochodami poniżej 650 zł na osobę, z prawdopodobieństwem 62%;
- XXII węzeł końcowy — jeśli w gospodarstwie domowym było dwoje dzieci na utrzymaniu, głowa gospodarstwa legitymowała się wykształceniem średnim zawodowym lub ogólnokształcącym, a gospodarstwo domowe zlokalizowane było w mieście z liczbą mieszkańców ponad 500 tys., to, z prawdopodobieństwem 71%, nie należało ono do gospodarstw z dochodami poniżej 650 zł na osobę;
- XXIII węzeł końcowy — jeśli w gospodarstwie domowym było dwoje dzieci na utrzymaniu, głowa gospodarstwa miała wykształcenie średnie zawodowe lub ogólnokształcące, a gospodarstwo domowe zlokalizowane było w mieście z liczbą mieszkańców poniżej 500 tys. lub na wsi, to, z prawdopodobieństwem 59%, należało do grupy gospodarstw z dochodami poniżej 650 zł na osobę;
- XXIV węzeł końcowy — jeśli w gospodarstwie domowym nie było dzieci na utrzymaniu, gospodarstwo było czteroosobowe, a wiek głowy gospodarstwa wynosił powyżej 40 lat, to gospodarstwo takie kwalifikowało się do grupy gospodarstw z dochodami powyżej 650 zł na osobę, z prawdopodobieństwem 58%;
- XXV węzeł końcowy — jeśli w gospodarstwie domowym nie było dzieci na utrzymaniu, gospodarstwo było pięcio i więcej osobowe, a wiek głowy gospodarstwa wynosił powyżej 40 lat, to gospodarstwo takie częściej należało do grupy gospodarstw z dochodami poniżej 650 zł na osobę, z prawdopodobieństwem 56%;
- XXVII węzeł końcowy — jeśli w gospodarstwie domowym było jedno dziecko na utrzymaniu, głowa gospodarstwa domowego reprezentowała zasadniczy zawodowy poziom wykształcenia i miała ponad 40 lat, to gospodar-

stwo to, z prawdopodobieństwem 55%, zaliczono do gospodarstw z dochodami powyżej 650 zł;

- XXVIII węzeł końcowy — jeśli w gospodarstwie domowym było jedno dziecko na utrzymaniu, głowa gospodarstwa domowego reprezentowała zasadniczy zawodowy poziom wykształcenia, miała poniżej 40 lat i gospodarstwo domowe zlokalizowane było w mieście powyżej 100 tys. mieszkańców, to gospodarstwo takie, z prawdopodobieństwem 54%, zaliczone zostało do gospodarstw z dochodami powyżej 650 zł;
- XXIX węzeł końcowy — jeśli w gospodarstwie domowym było jedno dziecko na utrzymaniu, głowa gospodarstwa domowego reprezentowała zasadniczy zawodowy poziom wykształcenia, jej wiek nie przekraczał 40 roku życia i gospodarstwo domowe zlokalizowane było w mieście poniżej 100 tys. mieszkańców lub na wsi, to, z prawdopodobieństwem 60%, gospodarstwo to zaliczono do gospodarstw z dochodami poniżej 650 zł.

**ZESTAWIENIE (1) WYNIKÓW BADANIA GRUPY GOSPODARSTW DOMOWYCH
Z DOCHODAMI DO DYSPOZYCJI PONIŻEJ/POWYŻEJ 650 ZŁ NA OSOBĘ MIESIĘCZNIE**

Numer węzła	Gałąź		Liczba gospodarstw domowych z dochodami na osobę miesięcznie		Przewidywa- na klasa przynależno- ści gospo- darstw do węzła	Suma przypadków w węźle	Udział w węźle gospodarstw domowych z dochodami na osobę miesięcznie	
	lewo- stronna	prawo- stronna					poniżej 650 zł	powyżej 650 zł
w %								
1	2	3	7574	9932	1	17506	43	57
2	4	5	3130	7883	1	11013	28	72
3	6	7	4444	2049	0	6493	68	32
4	8	9	2123	2571	1	4694	45	55
5	x	x	1007	5312	1	6319	16	84
6	10	11	4247	1387	0	5634	75	25
7	x	x	197	662	1	859	23	77
8	12	13	764	1649	1	2413	32	68
9	14	15	1359	922	0	2281	60	40
10	x	x	2983	591	0	3574	83	17
11	16	17	1264	796	0	2060	61	39
12	x	x	533	1405	1	1938	28	72
13	18	19	231	244	1	475	49	51
14	x	x	325	123	0	448	73	27
15	20	21	1034	799	0	1833	56	44
16	22	23	863	691	0	1554	56	44
17	x	x	401	105	0	506	79	21
18	x	x	38	15	0	53	72	28
19	24	25	193	229	1	422	46	54
20	26	27	383	396	1	779	49	51
21	x	x	651	403	0	1054	62	38
22	x	x	48	119	1	167	29	71
23	x	x	815	572	0	1387	59	41
24	x	x	127	178	1	305	42	58
25	x	x	66	51	0	117	56	44
26	28	29	196	171	0	367	53	47
27	x	x	187	225	1	412	45	55
28	x	x	83	97	1	180	46	54
29	x	x	113	74	0	187	60	40

Źródło: obliczenia własne na podstawie danych z publikacji *Budżety...* (2008).

Następnie obliczono macierz błędnych obserwacji, która zestawia klasy obserwowane (w kolumnach) i klasy przewidywane (w wierszach). W tym przypadku, spośród 7574 gospodarstw domowych, które pozostawały w grupie z dochodami poniżej 650 zł na osobę, sklasyfikowano poprawnie 5392 gospodarstwa, a 2182 błędnie (tabl. 3). Trafność predykcji dla tej klasy wyniosła 0,71, czyli 71%. Jeśli chodzi o predykcję gospodarstw domowych, które zaliczono do grupy z dochodami powyżej 650 zł na osobę, to na 9932 przypadki błędnie zostały sklasyfikowane 1934 gospodarstwa, trafność predykcji dla tej grupy wyniosła 0,81, czyli 81% zostało sklasyfikowanych poprawnie.

Dla całego modelu wartość błędu wyniosła 23,5%, co obrazuje równanie: $(2182+1934)/(5392+1934+2182+7998) = 4116/17506 = 0,2351$, czyli 23,5%.

TABL. 2. LICZBA BŁĘDNYCH I PRAWIDŁOWYCH KLASYFIKACJI

Wyszczególnienie	Gospodarstwa domowe z dochodami do dyspozycji na osobę miesięcznie	
	poniżej 650 zł	powyżej 650 zł
Klasyfikacje:		
prawidłowe	5392	7998
błędne	2182	1934

Źródło: jak przy zestawieniu (1).

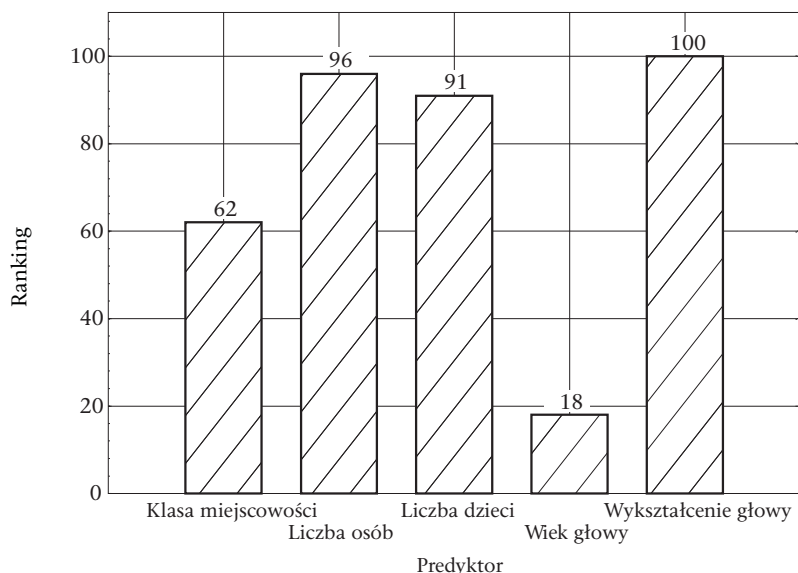
Ranking ważności poszczególnych predyktorów pokazał, że im większa wartość wskaźników, tym większy wpływ danej zmiennej na to, czy gospodarstwo domowe pozostaje czy też nie pozostaje w grupie z dochodami dyspozycyjnymi poniżej 650 zł na osobę miesięcznie. Największy względny wpływ na predykcję pozostawania w grupie gospodarstw domowych z dochodami powyżej 650 zł na osobę miała zmienna — wykształcenie głowy gospodarstwa domowego (wykr. 1). Na drugim miejscu znalazła się zmienna określająca wielkość gospodarstwa domowego, a na trzecim — liczba dzieci na utrzymaniu do 26 roku życia. Pewien wpływ wywierała także klasa miejscowości, a najmniejszy — wiek głowy gospodarstwa domowego.

Kolejna analiza dotyczyła grupy gospodarstw domowych pracowników, podzielonych na gospodarstwa z dochodami do dyspozycji równymi bądź powyżej 2 tys. zł na osobę i gospodarstwa domowe z dochodami poniżej 2 tys. zł na osobę. W tym przypadku wielkość drzewa, czyli liczba liści wyniosła 6, a głębokość — 5 (schemat 2). W zbudowanym modelu otrzymano 6 węzłów końcowych, za pomocą których można opisać analizowane gospodarstwa domowe.

Poszczególne liście można zinterpretować następująco:

- II węzeł końcowy — jeśli w gospodarstwie domowym głowa gospodarstwa miała wykształcenie podstawowe, zasadnicze zawodowe, średnie ogólnokształcące i zawodowe, to należało ono do grupy gospodarstw, w których średni dochód do dyspozycji wynosił poniżej 2 tys. zł na osobę miesięcznie, z prawdopodobieństwem 98%;

Wykr. 1. RANKING WAŻNOŚCI PREDYKATORÓW POZOSTAWANIA LUB NIE W GRUPIE GOSPODARSTW DOMOWYCH Z DOCHODAMI DO DYSPOZYCJI PONIŻEJ 650 ZŁ NA OSOBĘ MIESIĘCZNIE (0 — niska ważność, 100 — wysoka ważność)

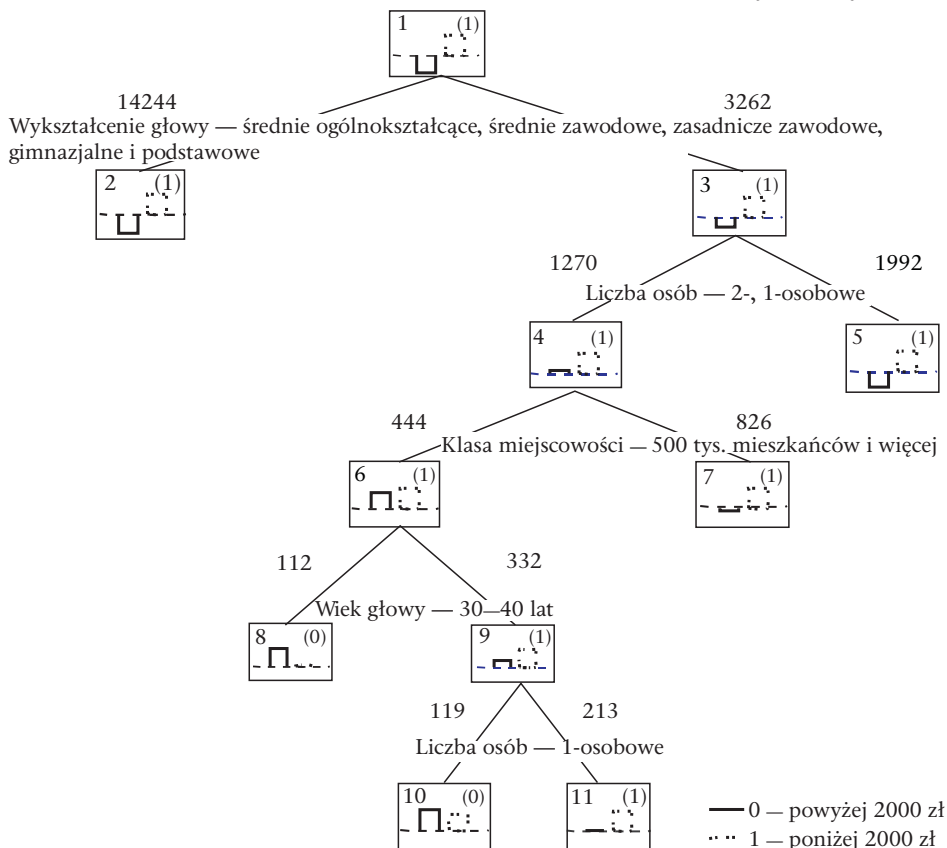


Źródło: jak przy wykr. 1.

- V węzeł końcowy — jeśli głowa gospodarstwa domowego miała wykształcenie wyższe, a liczba członków gospodarstwa domowego stanowiła ponad dwie osoby, to należało ono do grupy gospodarstw, w których średni dochód do dyspozycji wynosił poniżej 2 tys. zł na osobę miesięcznie, z prawdopodobieństwem 91% (wykr. 3, tabl. 5);
- VII węzeł końcowy — jeśli głowa gospodarstwa domowego miała wyższe wykształcenie, gospodarstwo domowe było jedno- lub dwuosobowe i było zlokalizowane w mieście z liczbą mieszkańców poniżej 500 tys. lub na wsi, to gospodarstwo takie należało do grupy gospodarstw, w których średni dochód do dyspozycji wynosił poniżej 2 tys. zł na osobę miesięcznie, z prawdopodobieństwem 69%;
- VIII węzeł końcowy — jeśli głowa gospodarstwa domowego miała wyższe wykształcenie, gospodarstwo domowe było jedno- lub dwuosobowe, było ono zlokalizowane w mieście z liczbą mieszkańców powyżej 500 tys. i wiek głowy gospodarstwa domowego wynosił od 30 do 40 lat, wówczas to gospodarstwo, z prawdopodobieństwem 66%, należało do grupy gospodarstw, w których średni dochód do dyspozycji wynosił powyżej 2 tys. zł na osobę miesięcznie;
- X węzeł końcowy — jeśli głowa gospodarstwa domowego miała wyższe wykształcenie, gospodarstwo domowe było jednoosobowe, zlokalizowane

w mieście z liczbą mieszkańców powyżej 500 tys., a wiek głowy gospodarstwa domowego wynosił mniej niż 30 lat lub więcej niż 40 lat, wówczas to gospodarstwo domowe, z prawdopodobieństwem 54%, należało do grupy gospodarstw, w których średni dochód do dyspozycji wynosił powyżej 2 tys. zł na osobę miesięcznie;

Schemat. 2. DRZEWO KLASYFIKACYJNE CART GOSPODARSTW DOMOWYCH PRACOWNIKÓW Z DOCHODAMI DO DYSPOZYCJI NA OSOBĘ POWYŻEJ (0) LUB PONIŻEJ (1) 2000 ZŁ NA OSOBĘ MIESIĘCZNIE



Źródło: jak przy wykr. 1.

- XI węzeł końcowy — jeśli głowa gospodarstwa domowego miała wyższe wykształcenie, gospodarstwo domowe było dwuosobowe, zlokalizowane w mieście z liczbą mieszkańców powyżej 500 tys., a wiek głowy gospodarstwa domowego wynosił poniżej 30 roku życia lub ponad 40 lat, to takie gospodarstwo należało do grupy gospodarstw, w których średni dochód do dys-

pozycji wynosił poniżej 2 tys. zł na osobę miesięcznie, z prawdopodobieństwem 67%.

ZESTAWIENIE (2) WYNIKÓW BADANIA GRUPY GOSPODARSTW DOMOWYCH Z DOCHODAMI DO DYSPOZYCJI PONIŻEJ/POWYŻEJ 2000 ZŁ NA OSOBĘ MIESIĘCZNIE

Numer węzła	Gałąź		Liczba gospodarstw domowych z dochodami na osobę miesięcznie		Przewidywa- na klasa przynależno- ści gospo- darstw do węzła	Suma przypadków w węźle	Udział gospodarstw domo- wych w węźle z dochodami na osobę miesięcznie	
	lewo- stronna	prawo- stronna					poniżej 2 tys.	powyżej 2 tys.
w %								
1	2	3	939	16567	1	17506	5	95
2	x	x	290	13954	1	14244	2	98
3	4	5	649	2613	1	3262	20	80
4	6	7	461	809	1	1270	36	64
5	x	x	188	1804	1	1992	9	91
6	8	9	209	235	1	444	47	53
7	x	x	252	574	1	826	31	69
8	x	x	74	38	0	112	66	34
9	10	11	135	197	1	332	41	59
10	x	x	64	55	0	119	54	46
11	x	x	71	142	1	213	33	67

Ź r ó d ł o: jak przy zestawieniu (1).

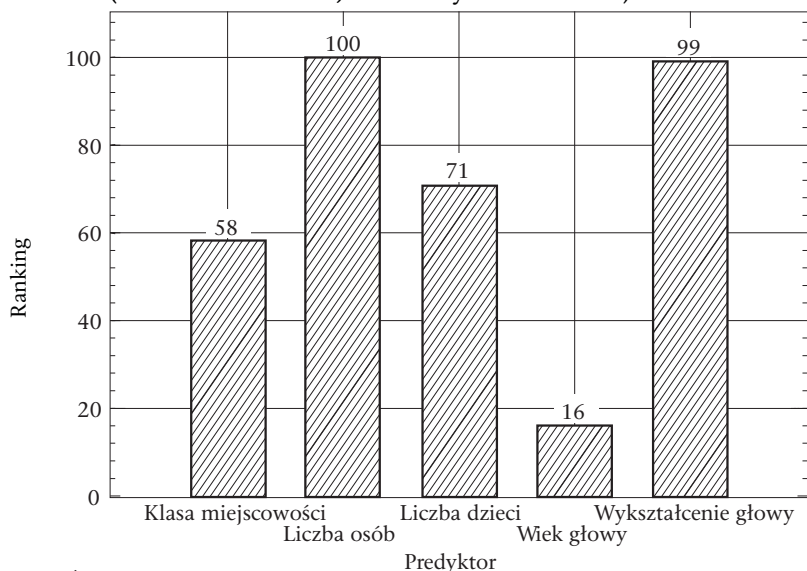
Zestawienie klas obserwowanych i klas przewidywanych wykazało, że spośród 939 gospodarstw domowych, które pozostawały w grupie z dochodami powyżej 2 tys. zł na osobę, prawidłowo sklasyfikowano 138 gospodarstw, a błędnie — 801. Trafność predykcji dla tej klasy wyniosła 0,17, czyli 17%. Jeśli chodzi o gospodarstwa domowe, które były zakwalifikowane do grupy z dochodami poniżej 2 tys. zł na osobę, to na 16567 przypadków 93 zostały błędnie sklasyfikowane, a prawidłowo — 16474. Dla całego modelu wartość błędu wyniosła 0,051, czyli 5,1%. Należy podkreślić, że o przydatności modelu nie decyduje wyłącznie trafność predykcji całego modelu, ale również trafność predykcji poszczególnych klas.

TABL. 3. LICZBA BŁĘDNYCH I PRAWIDŁOWYCH KLASYFIKACJI

Wyszczególnienie	Gospodarstwa domowe z dochodami do dyspozycji na osobę miesięcznie	
	powyżej 2 tys. zł	poniżej 2 tys. zł
Klasyfikacje:		
prawidłowe	138	16474
błędne	801	93

Ź r ó d ł o: jak przy zestawieniu (1).

**Wykr. 2. RANKING WAŻNOŚCI PREDYKATORÓW POZOSTAWANIA LUB NIE
W GRUPIE GOSPODARSTW DOMOWYCH Z DOCHODAMI
DO DYSPOZYCJI POWYŻEJ 2000 ZŁ NA OSOBĘ MIESIĘCZNIE
(0 — niska ważność, 100 — wysoka ważność)**



Źródło: jak przy wykr. 1.

Przyglądając się rankingowi ważności predyktorów zauważono, że największy względny wpływ na predykcję pozostawania w określonej grupie miała wielkość gospodarstwa domowego, prawie taki sam wpływ wywierała zmienna — wykształcenie głowy gospodarstwa domowego (wykr. 2). Na trzecim miejscu znalazła się zmienna — liczba dzieci na utrzymaniu do 26 roku życia, na czwartym — klasa miejscowości zamieszkania, a najmniejszy wpływ miał wiek głowy gospodarstwa domowego.

Podsumowując, w obu modelach zmiennymi w największym stopniu wpływającymi na predykcję pozostawania w określonej grupie były zmienne — wykształcenie głowy gospodarstwa domowego oraz wielkość gospodarstwa domowego. W grupie gospodarstw domowych pracowników największy udział miały gospodarstwa domowe, w których głowa gospodarstwa legitymowała się wykształceniem zasadniczym zawodowym (37%) oraz średnim zawodowym (25%) (tabl. 4). Najmniejszy udział stanowiły gospodarstwa domowe, w których głowa gospodarstwa reprezentowała średni ogólnokształcący poziom wykształcenia. Średni miesięczny dochód do dyspozycji kształtował się od 581 zł na osobę w gospodarstwach domowych z głową gospodarstwa reprezentującą podstawowy poziom wykształcenia do 1507 zł na osobę w gospodarstwach domowych z głową gospodarstwa z wyższym wykształceniem.

TABL. 4. STRUKTURA GOSPODARSTW DOMOWYCH ORAZ ŚREDNI MIESIĘCZNY DOCHÓD DO DYSPOZYCJI NA OSOBĘ WEDŁUG POZIOMU WYKSZTAŁCENIA GŁOWY GOSPODARSTWA DOMOWEGO W 2006 R.

Wykształcenie głowy gospodarstwa domowego	Udział gospodarstw domowych w badanej grupie gospodarstw w %	Miesięczny dochód do dyspozycji na osobę w zł
Gimnazjalne i podstawowe	10,2	581
Zasadnicze zawodowe	37,0	647
Średnie zawodowe	25,4	887
Średnie ogólnokształcące	8,7	981
Wyższe	18,6	1507

Ź r ó d ł o: jak przy zestawieniu (1).

Gospodarstwa domowe pracowników to głównie gospodarstwa trzy- i cztero-osobowe (56%), najmniejszy udział stanowiły jednoosobowe gospodarstwa domowe (ok. 7%). Średni dochód na osobę wynosił od 1701 zł miesięcznie w gospodarstwach jednoosobowych do 537 zł miesięcznie w gospodarstwach domowych pięć i więcej osobowych.

TABL. 5. STRUKTURA GOSPODARSTW DOMOWYCH ORAZ ŚREDNI MIESIĘCZNY DOCHÓD DO DYSPOZYCJI NA OSOBĘ WEDŁUG WIELKOŚCI GOSPODARSTWA DOMOWEGO W 2006 R.

Wielkość gospodarstwa domowego	Udział gospodarstw domowych w badanej grupie gospodarstw w %	Miesięczny dochód do dyspozycji gospodarstwa na osobę w zł
1-osobowe	6,7	1701
2	17,7	1251
3	27,1	905
4	28,7	707
5-osobowe i większe	19,8	537

Ź r ó d ł o: jak przy zestawieniu (1).

Podsumowanie wyników

Wielkość gospodarstwa domowego oraz poziom wykształcenia głowy gospodarstwa domowego to najważniejsze czynniki wpływające na predykcję pozostawania w określonej grupie ze względu na dochody w gospodarstwach domowych. Najmniejsze znaczenie wśród analizowanych zmiennych miał wiek głowy gospodarstwa domowego.

Zmiennymi związanymi z pozostawaniem w grupie gospodarstw domowych z dochodami do dyspozycji poniżej 650 zł na osobę miesięcznie były: niższy poziom wykształcenia głowy gospodarstwa domowego, większa liczba osób

w gospodarstwie domowym oraz większa liczba dzieci, a także lokalizacja gospodarstwa w najmniejszych miastach i na wsiach.

Znalezieniu się w grupie gospodarstw domowych z dochodami dyspozycyjnymi wynoszącymi co najmniej 2 tys. zł na osobę miesięcznie sprzyjało wyższe wykształcenie głowy gospodarstwa domowego oraz mała liczba osób w gospodarstwie domowym (gospodarstwa 1-, 2-osobowe).

Uzyskane wyniki wskazują na celowość stosowania metod drzewkowych w tego rodzaju badaniach. Metoda ta charakteryzuje się łatwością w użyciu i przejrzystością graficznego wyniku analizy — drzewa.

dr Marlena Piekut — *Politechnika Warszawska*

LITERATURA

Budżety Gospodarstw Domowych 2006 (2008), GUS

Kurowski P. (2002), *Koszty minimum socjalnego i minimum egzystencji — dotychczasowe podejście*, www.ipiss.com.pl/opracowania_min.html

Łapczyński M. (2003), *Drzewa klasyfikacyjne w badaniach satysfakcji i lojalności klientów*, w: *Analiza satysfakcji i lojalności klientów*, „Zastosowanie statystyki i data mining”, StatSoft, Kraków

Łapczyński M., Sagan A. (2008), *Metody statystyczne w marketingu i badaniach rynku*, StatSoft, Kraków

Minimum socjalne, www.ipiss.com.pl/opracowania_min.html

SUMMARY

The author's purpose was to separate relatively homogenous household groups included to a defined group regarding a level of disposable incomes per capita. The Classification and Regression Trees (CART) method was used to realize the aim. Individual data of the CSO's household budget survey for 2006 were the material of the analysis. The household size as well as the education level of the household head were the most important factors affecting the predilection for remaining in a defined group regarding a level of income in households. The age of the household head had a less importance among analysed variables. Such features as the lower education level of the household head, more persons in a household as well as the higher number of children and localisation in the smallest towns and villages are favorable for remaining in the household group with lower income. Higher education level of the household head and lower person number in a household occur more often in households of the better economic position.