

Skuteczność modeli predykcji bankructwa skonstruowanych dla różnie wyselekcjonowanych zmiennych diagnostycznych

Bernard Kokczyński; bernard.kokczynski@edu.uni.lodz.pl
Dorota Witkowska; dorota.witkowska@uni.lodz.pl



Cel badawczy

Celem badania jest **porównanie liniowych modeli dyskryminacyjnych** skonstruowanych na podstawie zmiennych diagnostycznych **wybranych z wykorzystaniem różnych metod ich selekcji** oraz wskazanie, **który z modeli predykcji bankructwa okazał się najlepszy** z punktu widzenia wielowymiarowej oceny trafności modelu.

Hipotezy badawcze

Hipoteza główna

Wybór metody selekcji zmiennych diagnostycznych istotnie wpływa na skuteczność klasyfikacyjną modeli predykcji bankructwa przedsiębiorstw.

Hipoteza pomocnicza

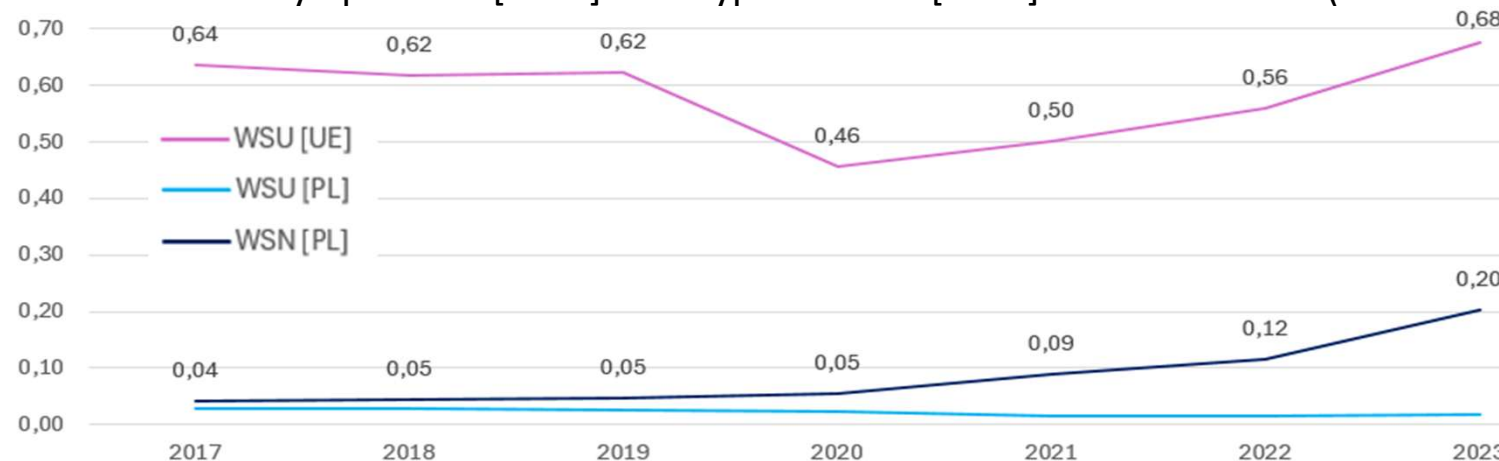
Statystyka λ Wilksa jest wiarygodnym wskaźnikiem efektywności klasyfikacyjnej modeli dyskryminacyjnych.

Uzasadnienie wyboru tematu

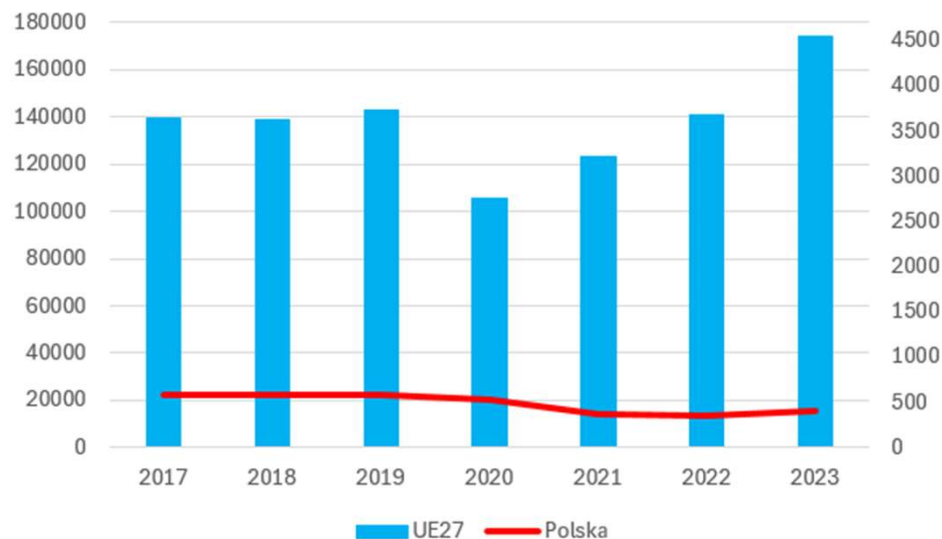
- W warunkach globalizacji, bankructwo firm może uruchomić **efekt domina**. Zjawisko to rozprzestrzenia się tym szybciej, im bardziej powiązana z innymi podmiotami jest firma dotknięta kryzysem (Mączyńska, 2013).
- Identyfikacja kryzysu w przedsiębiorstwie na **wczesnym etapie** pozwala **wdrożyć działania naprawcze lub szukać wsparcia finansowego**.
- Potrzeba **precyzyjnego** szacowania parametrów ryzyka kredytowego.
- Przedsiębiorstwa z sektora MŚP często **nie są wyposażone w mechanizmy wczesnego ostrzegania przed zagrożeniem upadłością**. Jednocześnie MŚP generuje około **45,3% wartości PKB** (PARP, 2024).

Zjawisko bankructwa w Polsce

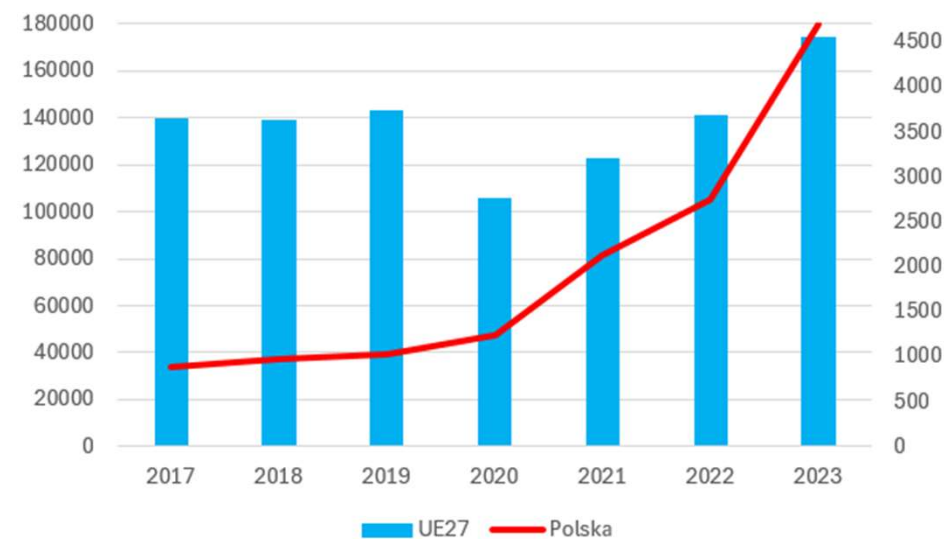
Wskaźniki struktury upadłości [WSU] i niewypłacalności [WSN] w UE27 i Polsce (2017-2023) [%]



Porównanie liczby sądowych upadłości przedsiębiorstw w krajach UE27 i Polsce w latach 2017-2023



Porównanie liczby sądowych upadłości przedsiębiorstw w krajach UE27 i łącznej liczby niewypłacalności w Polsce w latach 2017-2023



Linowe modele dyskryminacyjne w polskiej literaturze przedmiotu

Autor/rok	Równanie funkcji/Zmienne	Wartość graniczna	Autor/rok	Równanie funkcji/Zmienne	Wartość graniczna
Pogodzińska i Sojak (1995)	$Z_{PS} = 0,644741x_1 + 0,912304x_2$ x_1 - (aktywa obrotowe – zapasy) / zobowiązania krótkoterminowe; x_2 - wynik brutto / przychody ze sprzedaży	Punkt graniczny =0, Sfera pośrednia <- 0,454, 0,090>	Mączyńska i Zawadzki (2006)	$G = -1,498 + 9,498 \cdot \frac{WO}{A} + 3,566 \cdot \frac{KW}{A} + 2,903 \cdot \frac{WN + AM}{Z} + 0,452 \cdot \frac{MO}{ZKT}$ $\frac{WO}{A}$ - wynik operacyjny / aktywa ogółem; $\frac{KW}{A}$ - kapitał własny / aktywa ogółem; $\frac{WN+AM}{Z}$ - (wynik netto + amortyzacja) / zobowiązania ogółem; $\frac{MO}{ZKT}$ - aktywa obrotowe / zobowiązania krótkoterminowe.	Jeśli G> 0 to brak zagrożenia bankructwem
Gajdka i Stos (1996)	$Z_{GS} = 0,017803x_1 + 0,588694x_2 + 0,138657x_3 - 4,31026x_4 - 0,01038x_5 + 0,437449$ x_1 - wskaźnik bieżącej płynności; x_2 - wskaźnik zadłużenia aktywów; x_3 - wskaźnik rotacji aktywów w razach; x_4 - wskaźnik rentowności aktywów netto; x_5 - (zysk netto + odsetki) / przychody netto	Jeśli $Z_{GS} < 0,432589$ to brak zagrożenia bankructwem			
Hadasik (1998)	$Z_{HA4} = 2,36261 + 0,365425x_1 - 0,765526x_2 - 2,40435x_3 + 1,59079x_4 + 0,00230258x_5 - 0,0127826x_6$ x_1 - aktywa bieżące/zobowiązania bieżące; x_2 - (aktywa bieżące – zapasy)/zobowiązania bieżące; x_3 - zobowiązania ogółem/aktywa ogółem; x_4 - (aktywa bieżące – zobowiązania krótkoterminowe)/pasywa ogółem; x_5 - należności/przychody ze sprzedaży; x_6 - zapasy/przychody ze sprzedaży.	Jeśli $Z_{HA4}>0$ to brak zagrożenia bankructwem	Pociecha (2007)	$Y_1 = -1,843 - 0,486 \cdot ROE + 1,116 \cdot WKOP - 0,500 \cdot WPB$ ROE - wynik netto / kapitał własny; WKOP - (koszty działalności operacyjnej / przychody ze sprzedaży)·100%; WPB - aktywa obrotowe / zobowiązania krótkoterminowe.	Brak informacji
Wierzba (2000)	$Z_W = 3,26x_1 + 2,16x_2 + 0,3x_3 + 0,69x_4$ x_1 - (zysk z działalności operacyjnej – amortyzacja) / aktywa ogółem; x_2 - (zysk z działalności operacyjnej – amortyzacja) / sprzedaż produktów; x_3 - aktywa obrotowe/zobowiązania całkowite; x_4 - kapitał obrotowy/aktywa ogółem.	Jeśli $Z_W>0$ to brak zagrożenia bankructwem	Appenzeller i Szarzec (2007)	$D_{(w)} = 0,819 \cdot WP_2 + 2,567 \cdot WR_2 - 0,005 \cdot WS_2 + 0,006 \cdot WS_3 - 0,0095 \cdot WZ_2 - 0,556$ WP ₂ - majątek obrotowy/zobowiązania krótkoterminowe; WR ₂ - zysk operacyjny/przychody ze sprzedaży netto; WS ₂ - (średnia wartość zapasów/przychody netto ze sprzedaży) · liczba dni; WS ₃ - rotacja należności + rotacja zapasów; WZ ₂ - zobowiązania i rezerwy na zobowiązania/(wynik operacyjny + amortyzacja) · (12/okres obrachunkowy).	Wartość graniczna = 0
Model „Poznański” Hamrola, Czajka, Piechocki (2004)	$Z_{MP} = 3,562x_1 + 1,588x_2 + 4,288x_3 + 6,719x_4 - 2,368$ x_1 - wynik finansowy netto/majątek całkowity; x_2 - (majątek obrotowy – zapasy)/zobowiązania krótkoterminowe; x_3 - kapitał stały/majątek całkowity; x_4 - wynik finansowy ze sprzedaży/przychody ze sprzedaży.	Wartość graniczna = 0	Pociecha i inni (2014)	$D_7 = 0,5483 - 4,1702 \cdot R16 - 0,1071 \cdot R31$ R16 - (wynika na działalności operacyjnej + amortyzacja) / sumę bilansową; R31 - koszty operacyjne / zobowiązania krótkoterminowe.	Jeśli $D_7> 0,5$ to brak zagrożenia bankructwem
Hołda (2006)	$Z_H = 0,605 + (6,81 \cdot 10^{-1})PWP - (1,96 \cdot 10^{-2})SZ + (1,57 \cdot 10^1)RM + (9,69 \cdot 10^{-3})ZM + (6,72 \cdot 10^{-4})WOZ$ PWP - majątek obrotowy/zobowiązania krótkoterminowe; SZ - (zobowiązania ogółem/suma bilansowa)·100; ZM - [(zysk/strata netto)/średnioroczny majątek ogółem ²] · 100; RM - przychody z ogółu działalności/średnioroczny majątek ogółem; WOZ - (średnioroczne zobowiązania krótkoterminowe/koszt sprzedanych produktów, towarów i materiałów) · 360.	Jeśli $Z_H< -0,3$ zachodzi wysokie prawdopodobieństwo bankructwa; Sfera pośrednia <-0,3, 0,1>; Jeśli $Z_H> 0,1$ nie zachodzi prawdopodobieństwa ryzyka bankructwa	Kopczyński (2016)	$Z_k = 0,029x_1 + 1,311x_8 + 6,966x_{14} + 0,472x_{18} - 0,736$ x_1 - kapitał własny / kapitał obcy; x_8 - kapitał pracujący / aktywa ogółem; x_{14} - gotówka netto z działalności operacyjnej / przeciętny stan aktywów; x_{18} - przychody ze sprzedaży / przeciętny stan aktywów.	Jeśli $Z_k> 0$ to brak zagrożenia bankructwem
			Kopczyński (2022)	$Zk_{2022} = 0,176x_{13} + 0,33x_{18} + 2,493x_{30} + 0,068$ x_{13} - gotówka netto z działalności operacyjnej / zysk z działalności gospodarczej; x_{18} - przychody ze sprzedaży / przeciętny stan aktywów; x_{30} - (zysk netto·100%) / przeciętny stan aktywów.	Jeśli $Zk_{2022}> 0$ to brak zagrożenia bankructwem

Metody doboru zmiennych do modeli

- Arbitralny wybór własny na podstawie literatury przedmiotu oraz arbitralny wybór wg wskazań chatu GPT
- Metoda dwuetapowa: test istotności różnic między średnimi grupowymi oraz metoda krokowa postępująca
- Metoda Hellwiga doboru zmiennych diagnostycznych (centralne oraz izolowane)
- *Tstatystyka*
- Wsteczna metoda krokowa

Metoda dwuetapowa

- W pierwszym etapie identyfikowano zmienne najbardziej różnicujące obie grupy przedsiębiorstw (bankrutów i niebankrutów) w oparciu o zaimplementowane w pakiecie SPSS statystyki Lambda Wilksa i Fishera-Snedecora.
- W drugim etapie, konstruowano modele za pomocą metody krokowej postępującej. W tej procedurze doboru zmiennych na każdym etapie dodaje się do modelu zmienną, która najbardziej przyczynia się do poprawy skuteczności klasyfikacyjnej modelu, zaczynając od najważniejszej determinanty badanego zjawiska podobnie jak w pracach Pociecha i inni (2014) oraz Herman (2018).

Metoda dwuetapowa - statystyka Fishera-Snedecora

Sprawdzianem testu jest statystyka F o rozkładzie Fishera-Snedecora, wyrażająca stosunek wariancji międzygrupowej do wariancji wewnątrzgrupowej, co w przypadku dwóch grup można zapisać jako (Malarska, 2005):

$$F = \frac{\sum_{j=1}^2 (\bar{x}_j - \bar{x})^2 n_j}{\frac{1}{n-2} \sum_{i=1}^{n_j} \sum_{j=1}^2 (x_{ij} - \bar{x}_j)^2}$$

$\sum_{j=1}^2 (\bar{x}_j - \bar{x})^2 n_j$ to zmienność międzygrupowa wyrażająca różnicowanie obu średnich grupowych wokół średniej ogólnej,

$\sum_{i=1}^{n_j} \sum_{j=1}^2 (x_{ij} - \bar{x}_j)^2$ to zmienność wewnątrzgrupowa wyrażająca różnicowanie poszczególnych obserwacji wokół średnich grupowych,

$\bar{x}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_{ij}$ to średnie grupowe obliczone dla każdej z wyróżnionych grup ($j=1, 2$),

$\bar{x} = \frac{1}{n} \sum_{i=1}^{n_j} \sum_{j=1}^2 x_{ij}$ to średnia ogólna obliczona na podstawie wszystkich obserwacji,

n_j to liczba obserwacji w każdej grupie,

$n = \sum_{j=1}^2 n_j$ to liczba wszystkich obserwacji,

x_{ij} to obserwacja wyróżnionej cechy w i -tym obiekcie znajdującym się w j -tej grupie.

Metoda Hellwiga doboru zmiennych diagnostycznych

W analizie obiektów wielocechowych selekcja cech diagnostycznych opiera się na macierzy korelacji R , która służy do oceny ich przydatności:

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & r_{13} & \dots & r_{1K} \\ r_{21} & 1 & r_{23} & \dots & r_{2K} \\ \dots & \dots & \dots & \dots & \dots \\ r_{K1} & r_{K2} & r_{K3} & \dots & 1 \end{bmatrix}$$

dla każdej pary zmiennych x_i i x_j ($i, j = 1, 2, \dots, K$), r_{ij} to współczynnik korelacji liniowej Pearsona.

W macierzy korelacji $\mathbf{R}$, oblicza się sumę elementów każdej kolumny. Następnie wyszukuje się kolumnę p , dla której suma wartości bezwzględnych jest największa, a cecha odpowiadająca tej kolumnie jest uznawana za cechę centralną. Cechy odpowiadające wyróżnionym wierszom, są uznawane za cechy satelitarne. Następnie z macierzy korelacji R usuwa się wyróżnione kolumny i wiersze, co prowadzi do uzyskania zredukowanej macierzy korelacji – cechy izolowane.

Metoda *t*statystyka

Zgodnie z procedurą zaproponowaną przez Hermana (2018), dokonano wyboru 5 zmiennych, które charakteryzowały się najwyższą wartością bezwzględną statystyki t-Studenta, w teście porównującym średnią wartość wskaźników w badanych grupach, dla próbek niezależnych. Wartość statystyki testowej t oblicza się w przypadku prób o niejednorodnych wariancjach jako:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{N_1} + \frac{S_2^2}{N_2}}}$$

$\bar{x}_1; \bar{x}_2$ to średnie z prób, $S_1^2; S_2^2$ to wariancje z prób, $N_1; N_2$ to liczebności prób.

Statystyka ma rozkład t-Studenta o $\left(\frac{1}{2} + \frac{S_1^2 S_2^2}{S_1^2 + S_2^2}\right) (n_1 + n_2 - 2)$ stopniach swobody. Odrzucenie hipotezy zerowej następuje, jeśli $p\text{-value} < 0,05$. Podejście to uzupełniono o weryfikację redundancji informacyjnej, przeprowadzonej za pomocą macierzy korelacji Pearsona, gdzie za wartość krytyczną przyjęto 0,8.

Wsteczna metoda krokowa

Proces rozpoczyna się od pełnego zestawu zmiennych objaśniających, przyjętych na początku w modelu teoretycznym. Następnie w każdym kroku eliminowana jest zmienna najslabiej związana z badanym zjawiskiem, a proces ten kontynuowany jest, aż pozostaną jedynie istotne zmienne.

Raz usunięta zmienna nie jest ponownie uwzględniana w modelu w dalszych etapach analizy (Wiktorowicz i inni, 2020). W programie SPSS zmienne niezależne były usuwane z modelu na podstawie wartości statystyki Fishera-Snedecora.

Ewaluacja modeli

Ocena modeli była realizowana w oparciu o takie mierniki poprawności modeli jak: statystyka lambda Wilksa, AUC, skuteczność klasyfikacyjna bankrutów, niebankrutów oraz ogólna w próbie uczącej i testowej.

Sformułowano 2 wzorce – hipotetyczny i idealny. Wartości wzorcowe zostały wyznaczone jako:

- **hipotetyczne** - najlepsze wartości mierników w rozpatrywanym zbiorze modeli, czyli wartości minimalne dla statystyki lambda Wilksa oraz maksymalne dla mierników poprawności klasyfikacji.
- **idealne** - ekstremalne wartości wskazujące na doskonałą siłę dyskryminacyjną modelu lub 100% zdolność modelu do rozpoznawania obiektów,

Korzystając z 8 mierników oceniających modele predykcji bankructwa skonstruowano miernik syntetyczny bazujący na odległości euklidesowej od wzorca.

Ewaluacja modeli – odległość euklidesowa

Przed zastosowaniem miary odległości, należy dokonać normalizacji postaci zmiennych. Zastosowano standaryzację:

$$z_{ij} = (x_{ij} - \bar{x}_j) / S_j$$

x_{ij} oznacza wartość cechy, $\bar{x}_j$ to średnia wartość cechy, a S_j to wartość odchylenia standardowego cechy.

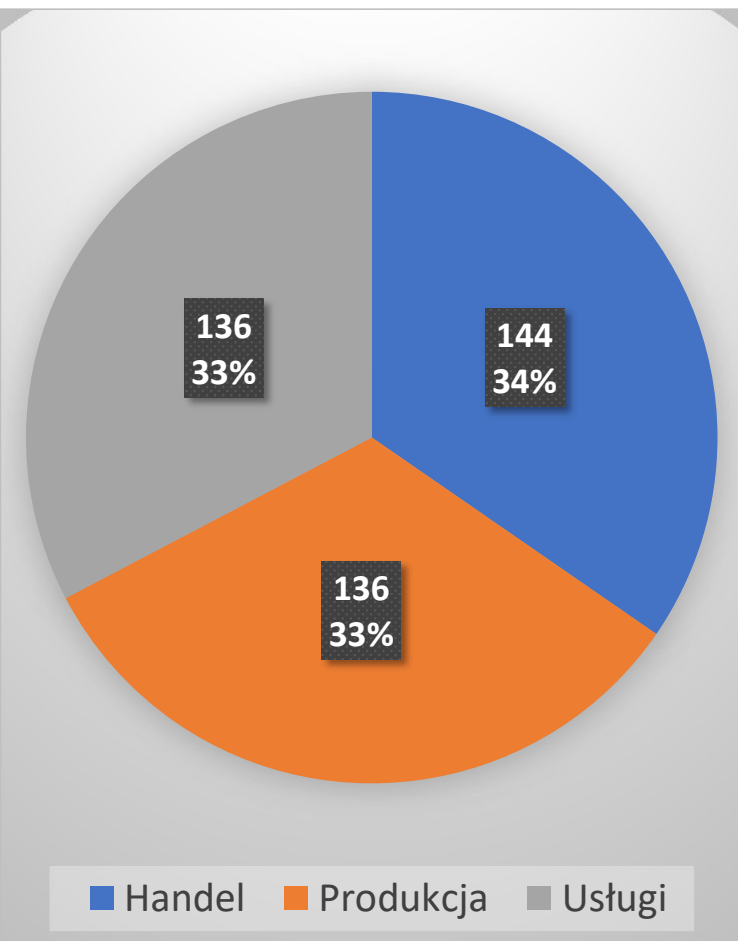
W ostatnim kroku policzono odległość euklidesową:

$$\left[\sum_{j=1}^n (z_{ij} - z_{0j})^2 \right]^{\frac{1}{2}}$$

n to liczba zmiennych, a z_{0j} to wzorcowa wartość po standaryzacji.

Zatem modelem uznanym za najlepszy zostanie ten, który będzie się charakteryzować najniższą wartością odległości euklidesowej od wzorca.

Baza danych - EMIS



Źródło: Opracowanie własne na podstawie danych dotyczących przedsiębiorstw pozyskanych z serwisu EMIS.

- Próba zawierała obserwacje dotyczące 416 przedsiębiorstw.
- Próba dobierana (choice-based sample) oraz próba dopasowana (matched sample).
- Bankrutów zdefiniowano jako firmy, które złożyły do sądu wnioski o upadłość.
- Każdemu bankrutowi odpowiadał podobny pod względem branży i wielkości rocznych przychodów niebankrut.
- Wnioski o upadłość zostały złożone przez przedsiębiorstwa w latach 2019-2022.
- W przypadku bankrutów uwzględniono sprawozdania z roku poprzedzającego złożenie wniosku do sądu o upadłość oraz sprzed dwóch lat. Identyczne podejście z zachowaniem tych samych lat badania zastosowano dla niebankrutów.

Zestawy zmiennych finansowych wybrane przy użyciu różnych metod selekcji

Metody doboru zmiennych	Symbol zestawu	Zestaw zmiennych
arbitralny wybór własny	A	P02, R04, S01, Z02
arbitralny wybór wg wskazań Chatu GPT	B	P02, R04, S07, W03, Z02
metoda dwustopniowa (wariant I)	C	R02, R09, W03
metoda dwustopniowa (wariant II)	D	R10, W03
Metoda Hellwiga (cechy centralne)	E	P02, R09, R10, R12, R13, S01, S04, S18, S19
Metoda Hellwiga (cechy izolowane)	F	D01, D02, R01, R02, R05, S03, S07, S09, W01, W02, W03, Z02, Z03, Z04
<i>tstatystyka</i>	G	R07, W02, W03, Z04, Z05
wsteczna metoda krokowa	H	R10, S20, W02, W03

Legenda: D – dynamika, P – płynność, R – rentowność, S – sprawność działania, W – wielkość przedsiębiorstwa, Z – zadłużenie.

Zestawy zmiennych finansowych wybrane przy użyciu różnych metod selekcji

R01 (EBITDA),
R02 (rentowność operacyjna do aktywów ogółem),
R04 (marża netto),
R05 (rentowność kapitału własnego),
R07 (zysk operacyjny do aktywów),
R09 (uśredniony zysk brutto do aktywów),
R10 (uśredniona rentowność aktywów obrotowych),
R12 (marża operacyjna po amortyzacji),
R13 (rentowność aktywów obrotowych),
W03 (logarytm aktywów),
W02 (logarytm struktury aktywów),
W01 (struktura aktywów),
D01 (dynamika przychodów),
D02 (dynamika kapitałów własnych),
P02 (wskaźnik płynności szybkiej),

S01 (wskaźnik rotacji aktywów),
S03 (zapasy do kosztów operacyjnych),
S04 (zapasy do przychodów ze sprzedaży),
S07 (konwersja należności),
S09 (rotacja zapasów w dniach),
S18 (uśredniony cykl konwersji gotówki netto),
S19 (uśrednione pokrycie kosztów operacyjnych zobowiązaniami krótkoterminowymi),
S20 (uśredniony wskaźnik rotacji aktywów),
Z02 (wskaźnik zadłużenia kapitału własnego),
Z03 (zadłużenie długoterminowe do kapitału własnego),
Z04 (udział kapitału własnego w sumie bilansowej),
Z05 (udział zobowiązań krótkoterminowych w sumie bilansowej)

Wyniki estymacji modeli

Funkcja dyskryminacyjna	Lambda Wilksa	AUC
A = 0,001·P02 + 0,007·Z02 + 0·R04 + 0,093·S01 -0,435	0,981	0,500
B = 0·P02 - 0,002·Z02 + 0·R04 + 0·S07 + 1,068·W03 -3,413	0,867	0,723
C = 0,217·R02 + 1,026·W03 + 0,032·R09 -3,322	0,870	0,722
D = 1,043·W03 + 0,003·R10 -3,357	0,863	0,730
E = 0·S18 + 0,271·R09 – 0,001·P02 + 0,028·R13 + 0·R14 – 0,018·R16 – 0,181 ·S04 + 0·S19 + 0,002·R10 + 0,243	0,952	0,715
F = -0,002·Z02 + 0,078·Z03 + 0·R01 + 0,129·R02 + 0·R05 – 0,061·S03 + 0·S07 – 0,051·W01 + 0,358·W02 + 0·D01 + 0,034·Z04 + 0,779·W03 – 0,001·D02 + 0 ·S09 - 2,133	0,798	0,763
G = 0,873·W03 + 0,074·Z04 + 0,286·W02 + 0,245·R07 + 0,045·Z05 - 2,606	0,856	0,718
H = 0,309·W02 + 0,926·W03 + 0,003·R10 + 0,081·S20 - 2,971	0,831	0,730

*Zerowe wartości niektórych parametrów modeli wynikają z zastosowanego zaokrąglania do trzeciego miejsca po przecinku.

Skuteczność klasyfikacyjna modeli

Model	Skuteczność klasyfikacyjna w próbie uczącej			Skuteczność klasyfikacyjna w próbie testowej		
	Bankruci	Niebankruci	Ogólna	Bankruci	Niebankruci	Ogólna
A	32,70%	76,20%	54,45%	37,50%	70,00%	53,75%
B	65,50%	72,60%	69,05%	70,00%	72,50%	71,25%
C	66,70%	77,40%	72,05%	70,00%	75,00%	72,50%
D	68,50%	77,40%	72,95%	70,00%	80,00%	75,00%
E	41,10%	91,10%	66,10%	42,50%	92,50%	67,50%
F	65,50%	76,80%	71,15%	72,50%	67,50%	70,00%
G	64,90%	73,80%	69,35%	67,50%	65,00%	66,25%
H	63,70%	79,20%	71,45%	67,50%	75,00%	71,25%

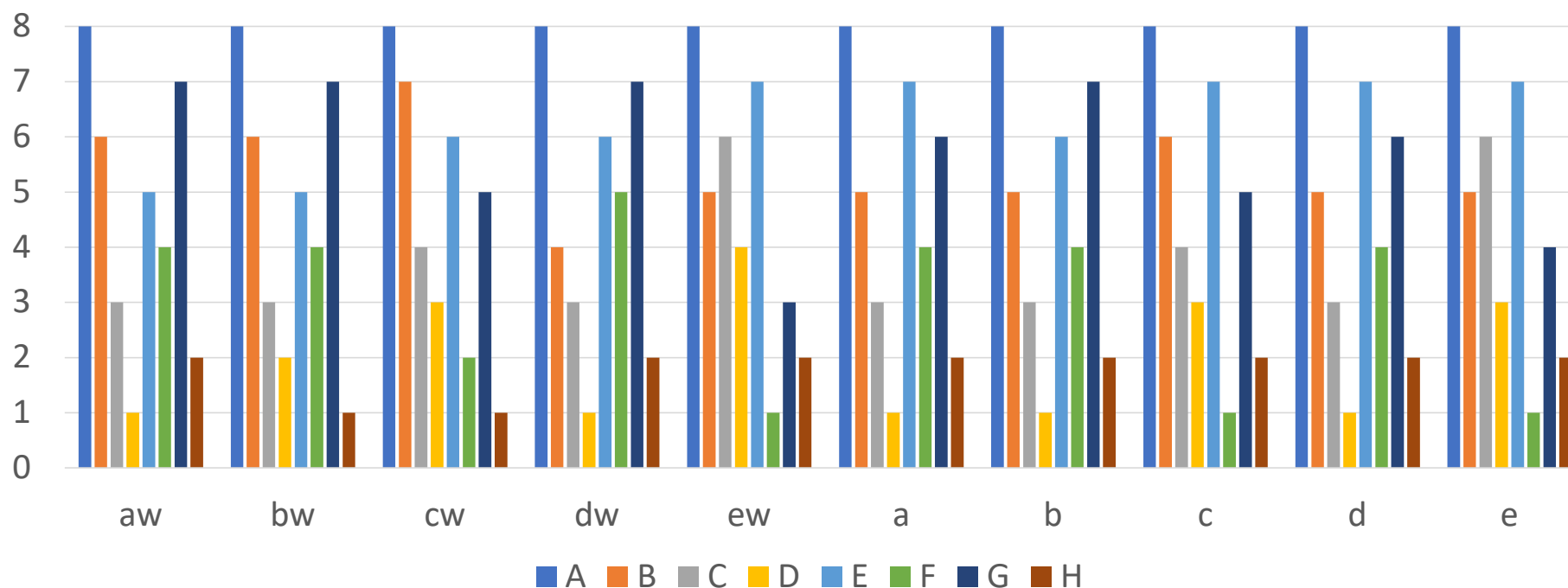
Proces wyłonienia najlepszej metody doboru zmiennych

W celu wyłonienia najlepszej metody doboru zmiennych diagnostycznych, policzono 10 odległości euklidesowych na podstawie:

- (a) wszystkich 8 mierników,
- (b) z wyłączeniem ogólnej klasyfikacji tj. dla 6 mierników,
- (c) z uwzględnieniem 4 mierników obejmujących jedynie lambda Wilksa, AUC oraz wyników klasyfikacji bankrutów i niebankrutów w próbie uczącej,
- (d) podobnie jak (c), ale dla próby testowej,
- (e) z uwzględnieniem 2 mierników - lambda Wilksa i AUC.

Warianty te porównano względem dwóch wzorców: modelu hipotetycznego (najlepszego w zbiorze) oraz modelu idealnego, co daje łącznie 10 odległości ($5 \cdot 2$ wzorce).

Porównanie modeli na podstawie odległości euklidesowej



Aby utworzyć ostateczny ranking wszystkich modeli, obliczono sumę rang utworzonych z pozycji osiągniętych przez poszczególne modele na podstawie przeprowadzonych analiz. W ten sposób otrzymano następującą kolejność modeli: H (z sumą rang 18), D (20), F (30), C (38), B (54), G (57), E (63) i A (80).

Uwaga:

symbole aw, bw, cw, dw i ew oznaczają odległość euklidesową od wzorca zdefiniowanego jako model hipotetyczny, tj. najlepszy spośród rozważanego zbioru modeli, natomiast a - e od wzorca będącego modelem idealnym.

Współczynniki korelacji rang Spearmana dla cech modeli dyskryminacyjnych

Nadane rangi wykorzystano do policzenia współczynnika korelacji Spearmana między wartością statystyki lambda Wilksa, a pozostałymi cechami modeli.

	Skuteczność klasyfikacyjna w próbie uczącej			Skuteczność klasyfikacyjna w próbie testowej			AUC
	Bankruci	Niebankruci	Ogólna	Bankruci	Niebankruci	Ogólna	
Współczynnik korelacji rang Spearmana	0,3631	-0,0417	0,5238	0,6131	-0,3512	0,2560	0,8393

Wnioski

- Modele H, D i F wydają się być najlepsze z punktu widzenia jednoczesnego uwzględnienia kilku miar dokładności. Twierdzimy zatem, że **metoda krokowa wsteczna, metoda dwuetapowa oraz metoda Hellwiga (zbiór zmiennych izolowanych)** okazały się **najlepszymi metodami doboru zmiennych diagnostycznych** do modeli predykcji bankructwa.
- Model F (zmienne zostały wybrane **metodą Hellwiga**) zawiera największą liczbę zmiennych, które opisują **wszystkie aspekty funkcjonowania przedsiębiorstwa**, co może przyczynić się do **większej stabilności w czasie** tego modelu (w porównaniu do modeli z małą liczbą zmiennych).
- **Najwyższą wartość odległości euklidesowej** uzyskał model A, tym samym arbitralny wybór okazał się **najmniej efektywną** metodą doboru zmiennych.

Wnioski

- W badaniu wartość statystyki lambda Wilksa **nie wydaje się dobrze wskazywać** na zdolności klasyfikacyjne modeli predykcji bankructwa, ponieważ **korelacja rang Spearmana** oceniana między rankingami sporządzonymi ze względu na wartości statystyki lambda Wilksa i według miar dokładności klasyfikacji była **raczej słaba**.
- Zatem **analiza skuteczności klasyfikacji** pozostaje **niezbędnym narzędziem** w ocenie jakości modeli dyskryminacyjnych.



Dziękuję za uwagę

Bernard Kokczyński; bernard.kokczynski@edu.uni.lodz.pl

Dorota Witkowska; dorota.witkowska@uni.lodz.pl

Skuteczność modeli predykcji bankructwa skonstruowanych dla różnie wyselekcjonowanych zmiennych diagnostycznych