

SERGIUSZ HERMAN

ANALIZA PORÓWNAWCZA ZDOLNOŚCI PREDYKCYJNEJ WYBRANYCH METOD PROGNOZOWANIA UPADŁOŚCI PRZEDSIĘBIORSTW

I. WSTĘP

Zagadnienie prognozowania upadłości przedsiębiorstw na świecie jest znane od początku XX w. Gwałtowny wzrost popularności tej problematyki przypada na lata dwudzieste i trzydzieste ubiegłego stulecia. Związany on był z panującym wówczas kryzysem gospodarczym. Prowadzone głównie w Stanach Zjednoczonych badania ukierunkowane były na wykorzystanie informacji płynących ze sprawozdań finansowych przedsiębiorstw w celu prognozowania ich upadłości. Pierwszymi, którzy podjęli taką próbę, byli w 1930 r. R.F. Smith oraz A.H. Winakor¹. Badali oni zmiany wartości 24 wskaźników finansowych, charakteryzujących sytuację 29 przedsiębiorstw zagrożonych upadłością. Na podstawie uzyskanych wyników wyselekcjonowali grupę 8 wskaźników mających dużą wartość prognostyczną. Dwa lata później podobną analizę przeprowadził P.J. Fitzpatrick². Tym razem wartości badanych wskaźników finansowych zostały porównane dla dwóch grup: 19 przedsiębiorstw sprawnie funkcjonujących oraz 19, wobec których ogłoszono upadłość. Na podstawie przeprowadzonego badania autor wskazał dwa wskaźniki (wynik finansowy/kapitał własny, kapitał własny/kapitał obcy), których wartości pozwalały najsprawniej przewidywać upadłość przedsiębiorstw.

Rozszerzeniem opisanych badań było opracowanie C.L. Merwina³. Na gruncie przeprowadzonej analizy porównawczej wyselekcjonował on trzy wskaźniki finansowe, tj. kapitał obrotowy/wartość aktywów, kapitał własny/kapitał obcy oraz aktywa bieżące/pasywa bieżące. Ich wartości średnie dla przedsiębiorstw sprawnie funkcjonujących oraz tych, wobec których ogłoszono upadłość, wyraźnie różniły się już na 6 lat przed zaistnieniem tej sytuacji. Przedstawiony kierunek badań kontynuował W.H. Beaver⁴. Badacz ten przeanalizował zdolność

¹ R.F. Smith, A.H. Winakor, *A test analysis of unsuccessful industrial companies*, „Bureau of Business Research” 1930, No. 31.

² P.J. Fitzpatrick, *A comparison of ratios of successful industrial enterprises with those of failed companies*, „The Certified Public Accountant” 1932, s. 598–605.

³ C. Merwin, *Financing small corporations in five manufacturing industries, 1926–1936*, National Bureau of Economic Research 1942.

⁴ W. Beaver, *Financial ratios as predictors of failure*, „Journal of Accounting Research” 5, 1966, s. 71–111.

predykcyjną 30 wskaźników finansowych. W swojej pracy zasygnalizował, że jednocześnie wykorzystanie wielu wskaźników powinno zagwarantować wyższą trafność niż prognoz opartych na zmiennych rozpatrywanych pojedynczo.

Krokiem milowym w rozwoju badań nad prognozowaniem upadłości przedsiębiorstw na świecie była publikacja z 1968 r. E.I. Altmana⁵. Jako pierwszy zaproponował on wykorzystanie wielowymiarowej analizy dyskryminacyjnej, wprowadzając tym samym niejako w życie pomysł Beavera (dotyczący jednoczesnego uwzględnienia wielu zmiennych przy prognozowaniu). Altman skonstruował narzędzie nazywane Z-score, wykorzystane do prognozowania upadłości przedsiębiorstw przemysłowych na podstawie pięciu wskaźników finansowych. Pionierska publikacja Altmana zapoczątkowała nowy – wielowymiarowy – kierunek badań. W kolejnych latach D. Martin⁶ zaproponował wykorzystanie w celu prognozowania upadłości przedsiębiorstw modeli regresji logistycznej, z kolei M. Żmijewski⁷ – modeli probitowych.

Wszystkie dotychczas przedstawione metody prognozowania upadłości przedsiębiorstw zaliczane są w literaturze do grupy metod statystycznych (ang. *statistical methods*). Ich główną wadą są liczne, związane z nimi założenia, takie jak np. rozkład normalny czy niezależność zmiennych predykcyjnych. Z biegiem lat coraz większą popularnością zaczęły się cieszyć, uwolnione od tychże restrykcyjnych założeń, metody zaliczane do grupy metod sztucznej inteligencji (ang. *artificial intelligent techniques*). Można tu wyróżnić m.in., wykorzystane po raz pierwszy przez M.D. Odoma i R. Sharda⁸, sieci neuronowe, algorytmy genetyczne⁹, metodę wektorów nośnych¹⁰ bądź też drzewa decyzyjne¹¹.

Częstość wykorzystania poszczególnych metod prognozowania upadłości przedsiębiorstw w badaniach empirycznych zmieniała się w czasie. Zbadali ją m.in. M.A. Aziz oraz H.A. Dar¹². Na podstawie ich opracowania można stwierdzić, że największą popularnością cieszą się metody z grupy statystycznych (pojawily się w 64% badanych publikacji), następnie miejsca zajmują, odpowiednio, metody miękkich technik obliczeniowych (25%) oraz modele teoretyczne (11%). Aziz i Dar przeanalizowali także szczegółowo częstość wykorzystania poszczególnych metod. Na wykresie 1 przedstawiono te, które znalazły najczęściej zastosowanie w badaniach.

⁵ E.I. Altman, *Financial ratios, discriminant analysis and the prediction of corporate bankruptcy*, „The Journal of Finance” 23(4), 1968, s. 589–609.

⁶ D. Martin, *Early warning of bank failures: a logit regression approach*, „Journal of Banking and Finance” 1, 1977, s. 249–276.

⁷ M. Żmijewski, *Methodological issues related to the estimation of financial distress prediction models*, „Journal of Accounting Research” 22, 1984.

⁸ M.D. Odom, R. Sharda, *A neural network model for bankruptcy prediction*, „Proceedings of IEEE International Conference on Neural Networks” 1990, s. 151–173.

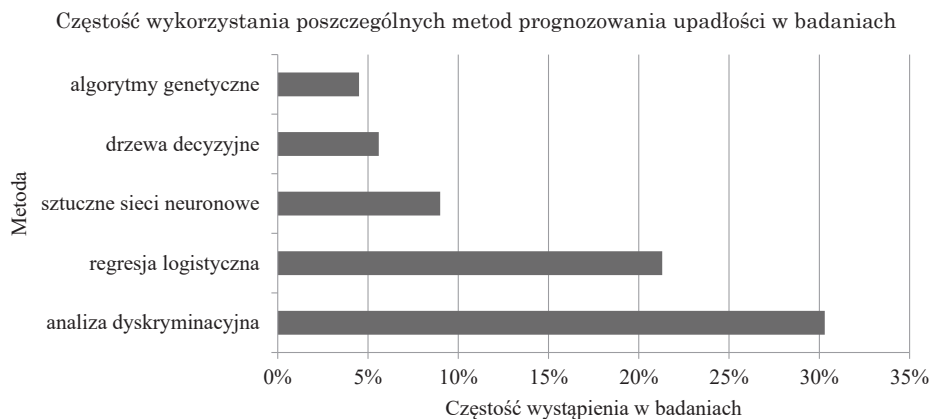
⁹ K.S. Shin, Y.J. Lee, *A genetic algorithm application in bankruptcy prediction modeling*, „Expert Systems with Applications” 23(3), 2002, s. 321–328.

¹⁰ K. Shin, T.S. Lee, H. Kim, *An application of support vector machines in bankruptcy prediction model*, „Expert Systems with Applications” 28, 2005, s. 127–135.

¹¹ A. Gepp, K. Kumar, *Predicting financial distress: a comparison of survival analysis and decision tree techniques*, „Procedia Computer Science” 54, 2015, s. 396–404.

¹² M.A. Aziz, H.A. Dar, *Predicting corporate bankruptcy – where we stand?*, „Corporate Governance Journal” 6(1), 2006, s. 18–33.

Wykres 1



Źródło: opracowanie własne na podstawie: M.A. Aziz, H.A. Dar, op. cit.

Na wyniki uzyskane przez Aziza oraz Dara duży wpływ ma okres, z którego pochodziły badane przez nich publikacje naukowe – wszystkie zostały udostępnione przed 2003 r.

Podobne badanie przeprowadził również E. Kirkos¹³. Tym razem przedmiotem analizy było 118 artykułów naukowych opublikowanych w okresie 2009–2011. Autor badania wskazał, iż w przypadku aż 55% badanych publikacji wykorzystano metody rozwinięte na bazie sztucznej inteligencji, co jest konsekwencją rozwoju technologii informacyjnych. Wśród nich najczęściej stosowanymi były metoda wektorów nośnych oraz sztuczne sieci neuronowe.

Autorzy badań przeprowadzanych za granicą wielokrotnie podejmowali próbę porównania zdolności prognostycznej (predykcyjnej) modeli konstruowanych za pomocą różnych metod prognozowania upadłości przedsiębiorstw. W literaturze obcojęzycznej można spotkać liczne opracowania będące zestawieniem tego typu badań: np. P. Ravi Kumar i V. Ravi¹⁴ dokonali kompleksowego przeglądu i charakterystyki aż 128 takich opracowań. W literaturze krajowej jest niewiele publikacji, których autorzy podejmują próbę porównania zdolności predykcyjnej różnych metod prognozowania upadłości.

Zasadniczym celem przedstawionego poniżej badania empirycznego jest kompleksowe porównanie zdolności predykcyjnej dziewięciu wybranych metod prognozowania upadłości przedsiębiorstw w Polsce. W tym celu została oszacowana zdolność prognostyczna szeregu modeli skonstruowanych z wykorzystaniem tychże metod. Zdolność ta mierzona jest w badaniu za pomocą globalnego współczynnika trafnych klasyfikacji (który określany jest ogólnie

¹³ E. Kirkos, *Assessing methodologies for intelligent bankruptcy prediction*, „Artificial Intelligence Review” 43(1), 2015, s. 83–123.

¹⁴ P. Ravi Kumar, V. Ravi, *Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A review*, „European Journal of Operational Research” 180, 2007, s. 1–28.

mianem trafności) oraz globalnego współczynnika błędnych klasyfikacji (określany dalej jako: błąd predykcji). Porównania zdolności predykcyjnej dokonano na podstawie wielu prób badawczych, przy uwzględnieniu różnych, statystycznych metod doboru zmiennych do modeli, różnej liczebności próby badawczej, a także zróżnicowanego horyzontu prognozy. Dotychczas badanie o tak szerokim zakresie nie zostało w Polsce przeprowadzone.

Artykuł został podzielony na trzy części. W pierwszej z nich przedstawiono przegląd krajowych badań empirycznych, których autorzy wykorzystali różne metody w celu prognozowania upadłości przedsiębiorstw. W kolejnej części artykułu scharakteryzowano próbę oraz metodykę badawczą. W ostatniej części publikacji przedstawiono wyniki przeprowadzonego badania empirycznego. Rozważania zamykają wnioski.

II. PRZEGLĄD LITERATURY

Jak wspomniano we wstępie, w literaturze krajowej jest niewiele publikacji naukowych, których autorzy podejmują próbę porównania zdolności predykcyjnej modeli skonstruowanych za pomocą różnych metod prognozowania upadłości przedsiębiorstw. Pierwszym, który zdecydował się na wykorzystanie w tym celu więcej niż dwóch metod statystycznych był A. Hołda¹⁵. Wykorzystana przez niego próba badawcza składała się z 374 przedsiębiorstw, reprezentujących trzy branże gospodarki, tj. produkcyjną, budowlaną oraz handlowo-usługową. Próbę tę podzielono na grupę uczącą oraz testową, zgodnie z proporcją 70:30. Hołda w swoim badaniu wykorzystał cztery metody prognozowania upadłości przedsiębiorstw, tj. liniową analizę dyskryminacyjną, regresję logistyczną, sieci neuronowe oraz drzewa decyzyjne, a także trzy różne metody doboru zmiennych. Uzyskane w badaniu wyniki pozwoliły na sformułowanie wniosku, że nie istnieje jedna metoda czy technika budowy modeli predykcji upadłości, która miałaby systematyczną przewagę nad pozostałymi.

Kolejnym autorem, który zdecydował się na wykorzystanie różnych metod prognostycznych w badaniach, był T. Korol. W tym celu analizował on grupę 185 spółek z sektora usług oraz produkcji, z czego 53 firmy stanowiły próbę uczącą. Autor badania wykorzystał dwie próby testowe, liczące odpowiednio 54 oraz 132 przedsiębiorstwa. W pierwszej analizie empirycznej Korol¹⁶ wykorzystał następujące metody prognostyczne:

- rekurencyjną sztuczną sieć neuronową (SSN REC),
- jednokierunkową wielowarstwową sztuczną sieć neuronową (SSN MLP),
- sztuczną sieć neuronową opartą na algorytmach genetycznych (SSN GA),
- sztuczną sieć neuronową o radialnych funkcjach bazowych (SSN RBF),

¹⁵ A. Hołda, *Zasada kontynuacji działalności i prognozowanie upadłości w polskich realiach gospodarczych*, Wyd. AE w Krakowie, Kraków 2006.

¹⁶ T. Korol, *Prognozowanie upadłości firm przy wykorzystaniu miękkich technik obliczeniowych*, „Finansowy Kwartalnik Internetowy e-Finanse” 2010, nr 1.

- mapę samoorganizującą się (SOM),
- model wektorów nośnych (SVM),
- model logiki rozmytej (FL).

Badanie to było próbą weryfikacji skuteczności szerokiego wachlarza metod zaliczanych do grupy miękkich technik obliczeniowych w prognozowaniu upadłości firm w Polsce. Wyniki analizy jednoznacznie wykazały, że w przypadku wykorzystania proporcji przedsiębiorstw zbliżonej do rzeczywistych uwarunkowań gospodarki model logiki rozmytej charakteryzuje się wyższą zdolnością predykcyjną niż pozostałe modele sztucznej inteligencji.

W drugim badaniu, będącym rozszerzeniem zaprezentowanej właśnie analizy, Korol¹⁷ do grupy porównywanych poprzednio metod dołączył liniową analizę dyskryminacyjną, metodę regresji logistycznej, drzewa klasyfikacyjne oraz las losowy. Podobnie jak wcześniej, w przypadku większej próby testowej modele logiki rozmytej najtrafniej prognozowały upadłość badanych firm. W przypadku mniejszej próby testowej uzyskane wyniki uzależnione były od horyzontu stawianej prognozy.

Autorami kolejnej publikacji, o której należy wspomnieć w ramach dokonywanego przeglądu literatury, są J. Pociecha, B. Pawełek, M. Baryła oraz S. Augustyn¹⁸. Dysponowali oni próbami badawczymi o liczebności od 118 do 246 firm. W badaniu podzielono je na próbę uczącą oraz testową w proporcjach 7:3 i 6:4. Analizę empiryczną przeprowadzono w trzech wariantach:

a) I wariant – prognozowanie upadłości z rocznym wyprzedzeniem (łącznie w jedną grupę przedsiębiorstwa, wobec których ogłoszono upadłość w latach 2007–2010);

b) II wariant – prognozowanie upadłości z 2-letnim wyprzedzeniem (łącznie w jedną grupę przedsiębiorstwa, wobec których ogłoszono upadłość w latach 2007–2010);

c) III wariant – prognozowanie upadłości z 2-letnim wyprzedzeniem (biorąc pod uwagę wyłącznie przedsiębiorstwa, wobec których ogłoszono upadłość w 2009 r.).

W badaniu wykorzystano dwie metody doboru próby badawczej: metodę parowania oraz metodę losowania niezależnego. Autorzy porównali zdolność predykcyjną modeli skonstruowanych za pomocą 4 metod prognozowania upadłości przedsiębiorstw: liniowej analizy dyskryminacyjnej, regresji logistycznej, drzewa decyzyjnego CART oraz sieci neuronowej. Biorąc pod uwagę wszystkie powyższe założenia przeprowadzonego badania uzyskano łącznie 12 rankingów porównywanych modeli. Tym, który najczęściej pojawiał się na pierwszym miejscu wśród zestawień, jest model wykorzystujący sieć neuronową. Aż w 11 przypadkach plasował się on na szycie utworzonych rankingów. Tylko w jednym przypadku trafniejsze prognozy dawał model skonstruowany za pomocą metody drzew klasyfikacyjnych. Najgorszymi zdolnościami

¹⁷ T. Korol, *Systemy ostrzegania przedsiębiorstw przed ryzykiem upadłości*, Oficyna, Warszawa 2010.

¹⁸ J. Pociecha et al., *Statystyczne metody prognozowania bankructwa w zmieniającej się koniunkturze gospodarczej*, Fundacja UE w Krakowie, Kraków 2014.

prognostycznymi w ramach dwóch pierwszych wariantów badania odznaczały się modele statystyczne. W trzecim wariancie badania metodą o najniższych zdolnościach predykcyjnych było drzewo klasyfikacyjne.

Ostatnim autorem, który zdecydował się zastosować w swoich badaniach, związanych z prognozowaniem upadłości przedsiębiorstw różne metody badawcze, był D. Gąska. W pierwszym swoim opracowaniu empirycznym¹⁹ wykorzystał on próbę badawczą składającą się z zaledwie 36 spółek akcyjnych. Ze względu na tak niewielki rozmiar badanej grupy przedsiębiorstw w badaniu wykorzystano następujące metody szacowania błędu predykcji konstruowanych modeli:

- metodę resubstytucji,
- metodę 5-krotnej walidacji krzyżowej,
- metodę n -krotnej walidacji krzyżowej,
- metodę bootstrap 0,632.

Gąska porównał w swojej analizie zdolność predykcyjną modeli konstruowanych zgodnie z metodą wektorów nośnych z wynikami uzyskanymi dla metod regresji logistycznej oraz jednowymiarowego modelu referencyjnego²⁰. Najniższe błędy predykcji (a tym samym najwyższą zdolność predykcyjną) uzyskano dla metod wektorów nośnych z funkcją jądrową Gaussa.

W kolejnym swoim badaniu Gąska²¹ zdecydował się rozszerzyć zakres przedstawionego poprzednio badania. Próbę badawczą stanowiła tym razem grupa 94 spółek akcyjnych, natomiast wachlarz porównywanych metod prezentował się następująco:

- regresja logistyczna,
- las losowy,
- k -najbliższych sąsiadów,
- metoda wektorów nośnych z funkcją jądrową Gaussa,
- metoda wektorów nośnych z wielomianową funkcją jądrową,
- metoda klasyfikatorów rozmytych maksymalnego marginesu z funkcją jądrową Gaussa. Błąd predykcji został oszacowany za pomocą 5-krotnej walidacji krzyżowej. Najwyższą zdolnością predykcyjną charakteryzowały się metoda lasów losowych oraz metoda klasyfikatorów rozmytych maksymalnego marginesu.

Cechą wspólną trzech pierwszych przedstawionych w tej części artykułu badań jest fakt, że ich autorzy w celu oszacowania zdolności predykcyjnej modeli skonstruowanych za pomocą badanych metod prognostycznych wykorzystali wyłącznie prostą metodę podziału. Wyniki licznych badań empirycznych przeprowadzonych na świecie jednoznacznie wskazują na to, że metoda ta charakteryzuje się bardzo wysoką zmiennością uzyskanych

¹⁹ D. Gąska, *Zastosowanie metody SVM do oceny ryzyka bankructwa i prognozowania upadłości przedsiębiorstw*, „Śląski Przegląd Statystyczny” 2013, nr 11, s. 289–310.

²⁰ Wybór na podstawie próby uczącej najsilniej dyskryminującej cechy, ustalenie dla niej wartości progowej i dokonanie na jej podstawie klasyfikacji obiektów z próby testowej.

²¹ D. Gąska, *Prognozowanie bankructwa za pomocą klasyfikatorów rozmytych realizujących ideę maksymalnego marginesu*, „Śląski Przegląd Statystyczny” 13(19), 2015, s. 71–88.

za jej pomocą rezultatów²². W literaturze podkreśla się, że z prostej metody podziału należy korzystać tylko wtedy, gdy dysponuje się dostatecznie szerokim zbiorem danych, które pozwolą na wyodrębnienie odpowiednio licznych, niezależnych zbiorów: treningowego i testowego²³. Jedynym, który zdecydował się w swoich badaniach na inną metodę szacowania błędu predykcji, był Gąska. Należy jednak zaznaczyć, że jego rozważania ukierunkowane były przede wszystkim na zaproponowanie nowego podejścia w prognozowaniu upadłości przedsiębiorstw opierającego się na idei metody wektorów nośnych. Wykorzystane w badaniach inne metody prognozowania były jedynie punktem odniesienia przeprowadzanych analiz. Wspólnym mianownikiem wszystkich przedstawionych badań empirycznych jest fakt, że zdolność predykcyjna modeli konstruowanych za pomocą różnych metod prognostycznych była szacowana wyłącznie raz, na podstawie jednej próby badawczej.

III. PRÓBA I METODYKA BADAWCZA

Badanie empiryczne wymagało właściwej próby badawczej, reprezentującej dwie rozłączne populacje: przedsiębiorstw o złej (oznaczane dalej jako: „chore”) oraz dobrej kondycji finansowej (oznaczane dalej jako: „zdrowe”). Kryterium decydującym o zaklasyfikowaniu przedsiębiorstw do pierwszej grupy był fakt ogłoszenia przez odpowiedni sąd jego upadłości. W celu wyselekcjonowania próby wykorzystano informacje zawarte w Internetowym Monitorze Sądowym i Gospodarczym. W ten sposób zgromadzono dane finansowe dotyczące 90 spółek akcyjnych (w tym: 30 spółek akcyjnych z branży budownictwo [PKD 41.10-43.99z]), 30 spółek akcyjnych z branży przetwórstwo przemysłowe [PKD 10.11-33.20z] oraz 30 spółek akcyjnych z branży handel hurtowy i detaliczny [PKD 46.11-47.99z]).

Do każdego z tych przedsiębiorstw została dobrana spółka akcyjna o dobrej kondycji finansowej. Za kryteria dopasowania poszczególnych par przyjęto: sektor, działalność główną oraz wielkość aktywów. Dane finansowe spółek, wobec których ogłoszono upadłość, pochodziły z ich sprawozdań finansowych dotyczących roku poprzedzającego ten, w którym złożono pierwszy wniosek o ogłoszenie upadłości oraz dwóch poprzednich lat. Pochodziły one z okresu 2000–2013. Sprawozdania finansowe dla spółek zdrowych pochodziły z tych samych lat. Źródłem danych były bazy firm Notoria Serwis i Bisnode Dun & Bradstreet oraz Monitor Polski B.

W badaniach empirycznych obliczono i wykorzystano 19 wskaźników finansowych charakteryzujących rentowność, płynność, strukturę kapitałowo-

²² Np. U.M. Braga-Neto, E.R. Dougherty, *Is cross-validation valid for small-sample microarray classification?*, „Bioinformatics” 20(3), 2004; J. Kim, *Estimating classification error rate: repeated cross-validation, repeated hold-out and bootstrap*, „Computational Statistics & Data Analysis” 53(11), 2009.

²³ Np. B.D. Ripley, *Pattern recognition and neural networks*, CUP, Cambridge 1996.

-majątkową oraz sprawność działania przedsiębiorstw (tab. 1). Ich wyboru dokonano na podstawie przeglądu literatury – są to wskaźniki pojawiające się najczęściej w modelach prognozowania upadłości. W wyborze kierowano się także dostępnością danych w sprawozdaniach finansowych spółek.

Tabela 1

Lista wskaźników finansowych wykorzystanych w badaniach

Wskaźnik	Formuła
ROA	zysk netto/ aktywa
ROE	zysk netto/ kapitały własne
ZB	zysk brutto/ aktywa
ZS	zysk ze sprzedaży/ przychody ze sprzedaży
MZ	zysk brutto/ przychody ze sprzedaży
MZ2	zysk netto/ przychody ze sprzedaży
MZO	zysk operacyjny/ przychody ze sprzedaży
KP	kapitał pracujący/ suma bilansowa
WBP	majątek krótkoterminowy/ zobowiązania krótkoterminowe
WSP	(majątek obrotowy-zapasy)/ zobowiązania krótkoterminowe
WPP	(majątek obrotowy-zapasy-należności)/ zobowiązania krótkoterminowe
ZO	zobowiązania ogółem / aktywa ogółem
ZD	zobowiązania długoterminowe / aktywa ogółem
KW	kapitał własny/ aktywa ogółem
KWZ	kapitał własny/ zobowiązania ogółem
RN	średnia wartość należności/ przychody ze sprzedaży netto*365
RZ	średnia wartość zapasów/ przychody ze sprzedaży netto*365
RZob	średnia wartość zobowiązań/ przychody ze sprzedaży*365
Rakt	średnia wartość aktywów/ przychody ze sprzedaży*365

Źródło: opracowania własne.

W celu uniknięcia skorelowania zmiennych opisujących obiekty przyjęto, że każdorazowo przed procesem uczenia usuwane są te zmienne, które są silnie skorelowane z pozostałymi (współczynnik korelacji Spearmana wyższy od 0,90)²⁴. Dodatkowo w celu wyselekcjonowania tych wskaźników finansowych,

²⁴ W przypadku gdy dwie zmienne są silnie skorelowane z dalszej analizy usuwana jest ta z nich, dla której średnia z wartości bezwzględnych współczynników korelacji między tą zmienną z pozostałymi jest wyższa.

które charakteryzują się najwyższą zdolnością dyskryminacyjną, w badaniu wykorzystano trzy popularne, statystyczne metody doboru zmiennych:

– wybór 5 zmiennych, które charakteryzowały się najwyższą wartością bezwzględną statystyki t w teście porównującym średnią wartość wskaźników w badanych grupach – metoda oznaczona dalej jako: *tstatystyka*;

– dobór zmiennych na podstawie siły zależności między poszczególnymi wskaźnikami finansowymi a zmienną y opisującą przynależność obiektu do konkretnej grupy. Wykorzystano te zmienne, dla których współczynnik korelacji Spearmana jest statystycznie istotny przy poziomie $\alpha = 0,05$ – metoda oznaczona dalej jako: *korelacje*;

– krokowy dobór zmiennych wprzód. W badaniu przyjęto, że do modelu w kolejnych krokach włączana jest ta zmienna, która w największym stopniu powoduje spadek statystyki λ Wilksa modelu²⁵. Ponieważ rozkład tej statystyki jest bardzo złożony, za warunek wprowadzenia zmiennej przyjęto poziom istotności równy 0,1 dla wartości F (będącej dobrą aproksymacją λ Wilksa). Metoda oznaczona dalej jako: *krokowa*.

Jak wspomniano we wstępie, celem przeprowadzonego badania empirycznego jest porównanie zdolności predykcyjnej różnych metod prognozowania upadłości przedsiębiorstw. Miarą zdolności predykcyjnej może być błąd predykcji. Wyniki wcześniejszych badań autora wskazują, że w przypadku prognozowania upadłości spółek akcyjnych w Polsce estymatory błędu predykcji uzyskane za pomocą metody *bootstrap* charakteryzują się najbardziej pożądanymi własnościami²⁶. W związku z tym w badaniu wykorzystano estymator błędu predykcji +.632 zaproponowany w 1997 r. przez B. Efrona i R. Tibshiraniego²⁷.

W przeprowadzonej analizie porównawczej wykorzystano dziewięć następujących metod prognozowania upadłości przedsiębiorstw:

- liniową funkcję dyskryminacyjną,
- kwadratową funkcję dyskryminacyjną,
- regresję logistyczną,
- naiwny klasyfikator bayesowski,
- metodę k -najbliższych sąsiadów,
- drzewa klasyfikacyjne,
- las losowy,
- sieci neuronowe,
- metodę wektorów nośnych.

²⁵ Statystyka λ Wilksa określa moc dyskryminacyjną modelu, może przyjmować wartości z przedziału od zera (doskonała moc dyskryminacyjna) do jedności (zupełny brak mocy dyskryminacyjnej).

²⁶ S. Herman, *Analiza porównawcza wybranych metod szacowania błędu predykcji klasyfikatora*, „Przegląd Statystyczny” 63, 2016, z. 4.

²⁷ B. Efron, R. Tibshirani, *Improvements on cross-validation: the .632 bootstrap method*, „Journal of the American Statistical Association” 92(438), 1997.

Ich szczegółowy opis można znaleźć w pracach np. J. Koronackiego oraz J. Ćwika czy publikacji M. Krzyśko, W. Wołyńskiego, T. Góreckiego oraz M. Skorzybuta²⁸.

Większość z wykorzystanych metod wymagała określenia wartości odpowiednich parametrów. Ostatecznie wykorzystano te wartości parametrów, dla których metoda charakteryzowała się najniższą wartością oszacowanego błędu predykcji. Zatem w przypadku metody najbliższych sąsiadów rozważano liczbę sąsiadów z przedziału $\langle 1;6 \rangle$, ostatecznie decydując, że $k = 5$. Dla drzew klasyfikacyjnych w badaniu rozważano minimalną liczebność obiektów w węźle (aby podział był dokonywany): 5, 10, 15, 20, oraz minimalną liczebność liści: 1, 3, 5, 10. Finalnie wykorzystano drzewa klasyfikacyjne, w których minimalna liczebność obiektów w węźle wynosiła 5, natomiast minimalna liczebność liścia równa jest 1. Miarą różnorodności klas w węźle był wskaźnik Giniego. W przeprowadzonej analizie porównywano lasy losowe zbudowane na podstawie 100, 200, 300, 400 oraz 500 drzew, gdzie ilość zmiennych w węźle mogła wahać się w przedziale $\langle 1;5 \rangle$. Przedstawione w dalszej części wyniki dotyczą lasów losowych zbudowanych na podstawie 100 drzew, gdzie w każdym węźle losowana była tylko jedna zmienna. W przypadku sieci neuronowych wykorzystano postać perceptronu wielowarstwowego. Sieci składały się z 3 warstw (1 wejściowa, 1 ukryta oraz 1 wyjściowa). Liczba neuronów warstwy wejściowej odpowiadała liczbie wskaźników finansowych wykorzystanych do budowy poszczególnych sieci. Liczba neuronów warstwy ukrytej wynosiła 3. Uczenie sieci odbywało się z wykorzystaniem algorytmu BFGS. W przeprowadzonej analizie rozważano także różne warianty metody wektorów nośnych:

- z liniową funkcją jądrową;
- z funkcją jądrową Gaussa, z wartością parametru $\delta \in \langle 0; 0,9 \rangle$;
- z funkcją jądrową wielomianową, ze stopniem wielomianu $q \in \{1, 2, 3, 4, 5\}$;

Najwyższą trafnością prognoz charakteryzowała się metoda wektorów nośnych z liniową funkcją jądrową i ta została wykorzystana w przedstawionej w następnej części artykułu analizie empirycznej. Całość obliczeń została wykonana z wykorzystaniem środowiska statystycznego *R*.

IV. WYNIKI BADANIA

Badanie empiryczne przeprowadzono w trzech etapach. Na pierwszym zweryfikowano, w jaki sposób kształtuje się trafność klasyfikacji oraz utworzony na jej podstawie ranking badanych metod prognozowania upadłości przedsiębiorstw w zależności od wykorzystanej metody doboru zmiennych prognostycznych. W tym celu przyjęto następujące założenia:

²⁸ J. Koronacki, J. Ćwik, *Statystyczne systemy uczące się*, Wyd. Naukowo-Techniczne, Warszawa 2005; M. Krzyśko et al., *Systemy uczące się: rozpoznawanie wzorców, analiza skupień i redukcja wymiarowości*, Wyd. Naukowo-Techniczne, Warszawa 2008.

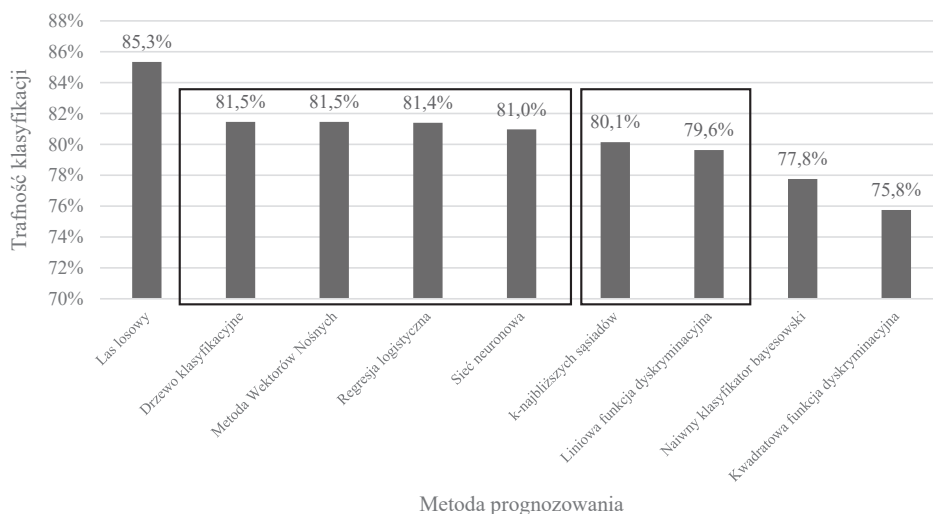
- a) błąd predykcji oszacowano dla 100 wylosowanych, stratyfikowanych próbek o liczebności 100 spółek,
- b) horyzont stawianej prognozy wynosił jeden rok,
- c) wykorzystano trzy przedstawione w trzeciej części artykułu metody doboru zmiennych, tzn. metody: *tstatystyka*, *korelacje*, *krokowa*.

Po oszacowaniu błędów predykcji modeli dla 100 wylosowanych próbek, z wykorzystaniem każdej z analizowanych metod prognozowania upadłości przedsiębiorstw, wyniki zostały uśrednione. W celu zweryfikowania, czy przeciętna trafność klasyfikacji jest równa dla analizowanych metod prognozowania upadłości przedsiębiorstw, wykorzystano jednoczynnikową analizę wariancji. By ustalić, które pary metod prognozowania upadłości przedsiębiorstw charakteryzują się odmiennym poziomem trafności klasyfikacji, wykorzystano test Tukeya.

Na wykresach 1–3 przedstawiono, w jaki sposób kształtowała się trafność klasyfikacji dla poszczególnych, analizowanych metod prognozowania upadłości przedsiębiorstw. Po przeprowadzeniu jednoczynnikowej analizy wariancji na podstawie wartości statystyk F stwierdzono, że we wszystkich trzech przypadkach (dla trzech metod doboru zmiennych predykcyjnych) są podstawy do odrzucenia hipotezy, że poziom trafności klasyfikacji dla wszystkich analizowanych metod prognostycznych jest jednakowy. W kolejnym kroku przeprowadzono test Tukeya. W przypadku metod zaznaczonych na wykresach ramkami nie ma podstaw do odrzucenia hipotezy, że przeciętny poziom trafności klasyfikacji jest dla nich równy.

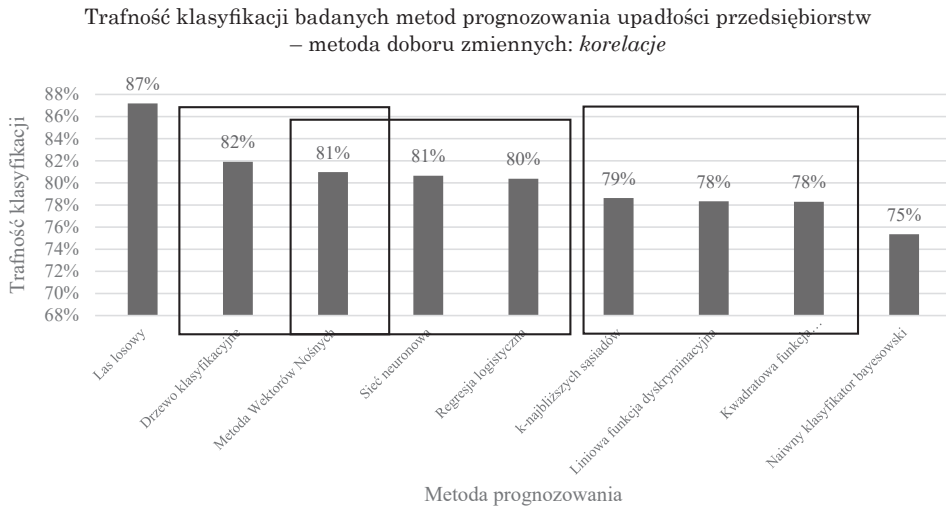
Wykres 1

Trafność klasyfikacji badanych metod prognozowania upadłości przedsiębiorstw – metoda doboru zmiennych: *tstatystyka*



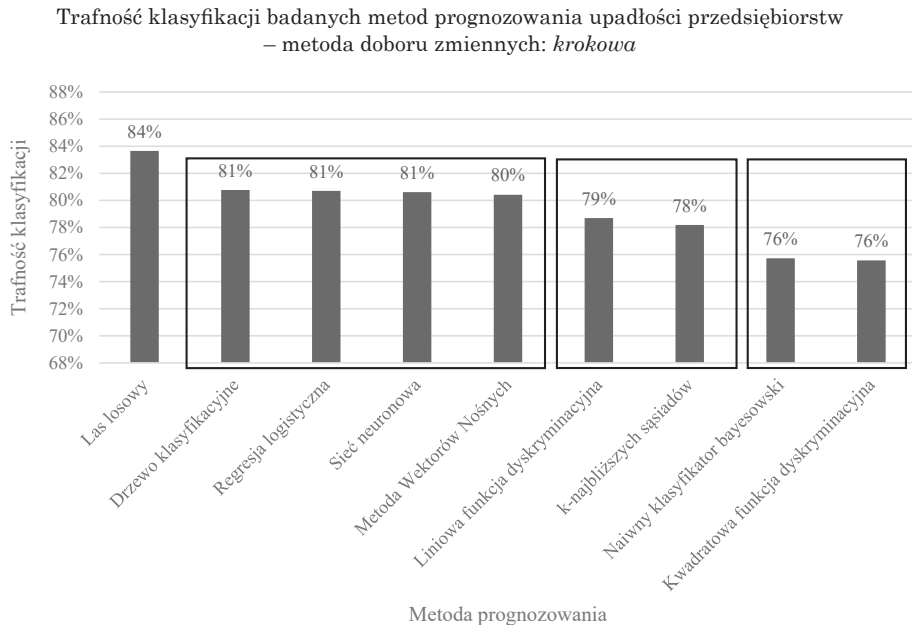
Źródło: opracowania własne.

Wykres 2



Źródło: opracowania własne.

Wykres 3



Źródło: opracowania własne.

Szczegółowa analiza wykresów dostarcza kilku interesujących wniosków. Przeciętny poziom trafności klasyfikacji dla analizowanych metod doboru zmiennych prognostycznych cechuje niewielka zmienność. Bez względu na

metodę doboru zmiennych do modelu metoda lasów losowych charakteryzuje się zawsze najwyższym, statystycznie istotnie różnym od innych metod, poziomem trafności klasyfikacji. Najmniej sprawnymi metodami prognozowania upadłości przedsiębiorstw okazały się z kolei naiwny klasyfikator bayesowski oraz kwadratowa funkcja dyskryminacyjna. Zmiana statystycznej metody doboru zmiennych predykcyjnych do modelu nie powoduje znaczących zmian w trafności klasyfikacji badanych metod, a tym samym ich rankingach.

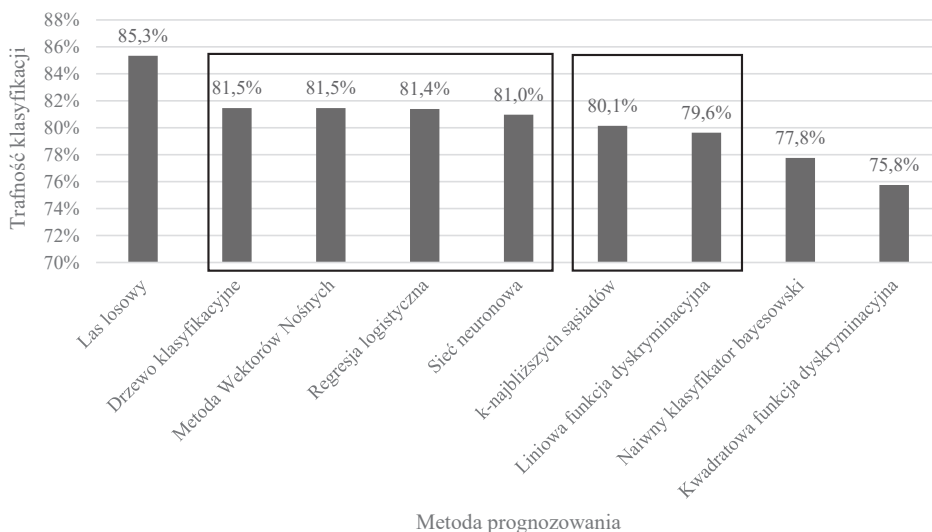
Na drugim etapie badania podjęto próbę zweryfikowania, jak kształtuje się trafność klasyfikacji modeli konstruowanych z wykorzystaniem badanych metod prognostycznych w zależności od liczebności próby badawczej. W tym celu:

- a) wykorzystano metodę doboru zmiennych prognostycznych: *tstatystyka*,
- b) przyjęto horyzont stawianej prognozy równy jeden rok,
- c) błąd predykcji oszacowano dla 100 wylosowanych prób o liczebnościach: 40, 70 i 100 spółek.

Na wykresach 4–6 przedstawiono przeciętny poziom trafności klasyfikacji dla badanych metod prognozowania upadłości przedsiębiorstw, przy uwzględnieniu różnej wielkości wylosowanych prób. Ponownie w celu zweryfikowania, czy ten przeciętny poziom jest jednakowy dla wszystkich metod, wykorzystano jednoczynnikową analizę wariancji. Uzyskane wyniki testu świadczą o tym, że są podstawy do odrzucenia tak sformułowanej hipotezy. Powtórnie zastosowano test Tukeya, w celu określenia, które metody prognozowania upadłości przedsiębiorstw charakteryzuje odmienna trafność klasyfikacji. Tak jak na pierwszym etapie badania metody, w których ta różnica nie jest statystycznie istotna, zostały zaznaczone ramkami.

Wykres 4

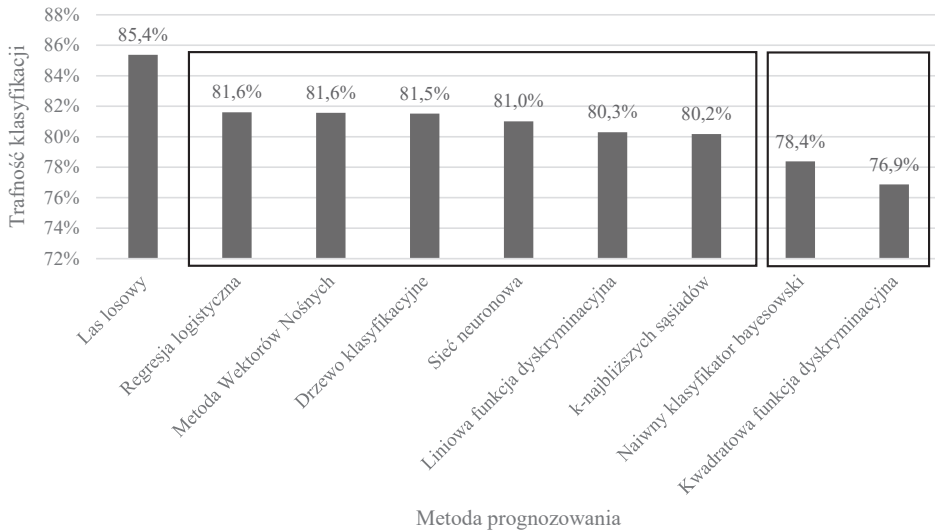
Trafność klasyfikacji badanych metod prognozowania upadłości przedsiębiorstw – wielkość wylosowanych prób: 100 spółek



Źródło: opracowania własne

Wykres 5

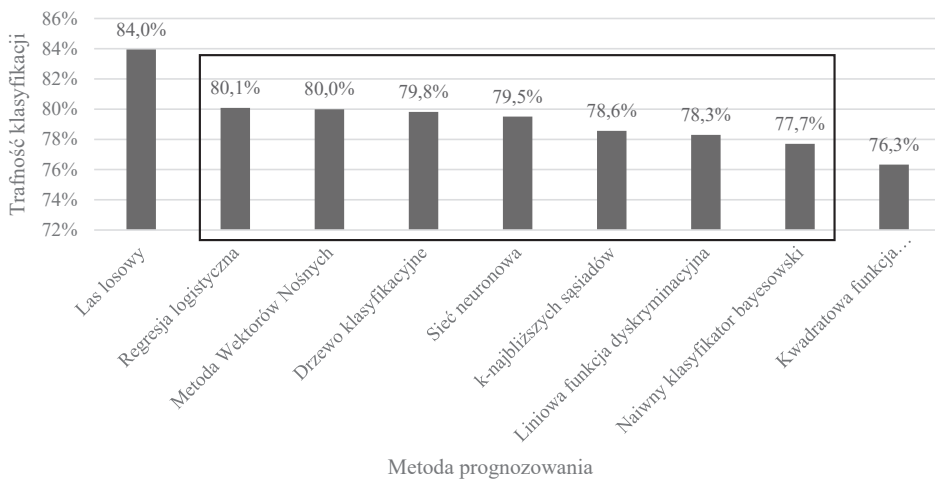
Trafność klasyfikacji badanych metod prognozowania upadłości przedsiębiorstw
– wielkość wylosowanych prób: 70 spółek



Źródło: opracowania własne.

Wykres 6

Trafność klasyfikacji badanych metod prognozowania upadłości przedsiębiorstw
– wielkość wylosowanych prób: 40 spółek



Źródło: opracowania własne.

Analiza wykresów pozwala stwierdzić, że niezależnie od wielkości próby badawczej metoda lasów losowych jest tą, która pozwala uzyskać najwyższą trafność klasyfikacji spośród wszystkich analizowanych metod. Najmniej sprawnymi metodami – charakteryzującymi się najniższą przeciętną trafnością klasyfikacji – okazały się za każdym razem kwadratowa funkcja dyskryminacyjna oraz naiwny klasyfikator bayesowski. Analiza wyników testu Tukeya (których odzwierciedleniem są ramki na wykresach) pozwala stwierdzić, iż w miarę zmniejszania się rozmiaru próby badawczej, różnice między trafnością klasyfikacji badanych metod zmniejszają się. W skrajnym przypadku – dla prób badawczych o liczebności 40 spółek – nie występuje statystycznie istotna różnica między przeciętnym poziomem trafności klasyfikacji aż dla 7 z 9 badanych metod prognozowania upadłości przedsiębiorstw. Na podstawie uzyskanych wyników można także wyciągnąć wniosek, że trafność klasyfikacji wszystkich badanych metod jest najmniejsza w przypadku najmniejszych prób badawczych (o liczebności 40 spółek).

Na ostatnim, trzecim etapie analizy zbadano, czy trafność klasyfikacji, a tym samym utworzony na jej podstawie ranking metod prognozowania upadłości przedsiębiorstw, uzależniony jest od okresu, na który stawiana jest prognoza. W tym celu przyjęto następujące założenia:

a) wykorzystano metodę doboru zmiennych prognostycznych: *tstatystyka*,

b) błąd predykcji oszacowano dla 100 wylosowanych podprób o liczebności 100 spółek,

c) przyjęto różne horyzonty prognoz. W tym celu wykorzystano wskaźniki finansowe charakteryzujące kondycję finansową przedsiębiorstw na:

- rok przed złożeniem wniosku o ogłoszenie upadłości przedsiębiorstw,
- dwa lata przed złożeniem wniosku o ogłoszenie upadłości przedsiębiorstw,
- trzy lata przed złożeniem wniosku o ogłoszenie upadłości przedsiębiorstw.

W tabeli 2 przedstawiono, w jaki sposób kształtuje się przeciętny poziom współczynników trafności klasyfikacji badanych metod, z uwzględnieniem różnej długości horyzontu stawianej prognozy. Zaprezentowane wyniki pozwalają stwierdzić, że w miarę wydłużania się okresu, na który stawiana jest prognoza, przeciętna trafność klasyfikacji wszystkich badanych metod się zmniejsza. Najmniejszy spadek tej zdolności predykcyjnej stwierdzono w przypadku liniowej analizy dyskryminacyjnej (spadek o 15,5 punktu procentowego). Analiza miejsc rankingowych poszczególnych metod badawczych pozwala stwierdzić, że trafność klasyfikacji w dużym stopniu uzależniona jest od horyzontu stawianej prognozy.

Tabela 2

Trafność klasyfikacji badanych metod prognozowania upadłości przedsiębiorstw w zależności od długości horyzontu prognozy

Metoda prognozowania	Horyzont prognozy					
	Rok		Dwa lata		Trzy lata	
	Trafność klasyfikacji	Ranking	Trafność klasyfikacji	Ranking	Trafność klasyfikacji	Ranking
Las losowy	85,3%	1	72,8%	1	63,9%	4
Drzewo klasyfikacyjne	81,5%	2	68,3%	5	59,6%	8
Metoda wektorów nośnych	81,5%	3	69,9%	3	64,1%	2
Regresja logistyczna	81,4%	4	70,7%	2	64,2%	1
Sieć neuronowa	81,0%	5	67,6%	6	60,5%	6
<i>K</i> -najbliższych sąsiadów	80,1%	6	66,1%	8	60,1%	7
Liniowa funkcja dyskryminacyjna	79,6%	7	69,8%	4	64,1%	3
Naiwny klasyfikator bayesowski	77,8%	8	66,5%	7	61,7%	5
Kwadratowa funkcja dyskryminacyjna	75,8%	9	63,7%	9	58,7%	9

Źródło: opracowania własne.

Wyłącznie metoda lasów losowych, niezależnie od długości okresu, na który stawiana jest prognoza, zawsze znajduje się wśród najsprawniejszych technik używanych w tym celu²⁹. W miarę wydłużania okresu, na który stawiana jest prognoza, coraz wyższe miejsce w rankingu badanych metod zajmuje, ciesząc się ogromną popularnością wśród praktyków, liniowa analiza dyskryminacyjna.

V. PODSUMOWANIE

W artykule dokonano przeglądu polskich badań empirycznych, których autorzy podjęli próbę wykorzystania różnych metod prognozowania upadłości przedsiębiorstw. Ze względu na ich wąski wachlarz, a także przyjętą metody-

²⁹ Zgodnie z wynikami testu Tukeya nie ma statystycznie istotnej różnicy między trafnością klasyfikacji czterech najsprawniejszych metod prognozowania upadłości przedsiębiorstw w przypadku 3-letniego horyzontu prognozy.

kę badawczą, podjęto próbę dokonania kompleksowej analizy porównawczej różnych metod prognozowania upadłości przedsiębiorstw. Metody zostały porównane ze względu na poziom trafności klasyfikacji, jaki można uzyskać za pomocą modeli skonstruowanych z ich wykorzystaniem.

Wyniki badania empirycznego dostarczyły kilku interesujących wniosków. Po pierwsze ranking badanych metod prognozowania upadłości przedsiębiorstw nie jest uzależniony od wykorzystanej metody doboru zmiennych predykcyjnych. Inaczej wygląda sytuacja w przypadku wielkości próby badawczej. Uzyskane rezultaty świadczą o tym, że rozmiar badanej grupy przedsiębiorstw ma wpływ na trafność klasyfikacji modeli skonstruowanych na podstawie różnych metod prognostycznych. W przypadku małych prób badawczych różnice między zdolnością predykcyjną odmiennych metod prognozowania upadłości przedsiębiorstw przestają być statystycznie istotnie różne. Wyniki analizy empirycznej wyraźnie wskazują również, że na zdolność predykcyjną analizowanych metod duży wpływ ma długość okresu, na który stawiana jest prognoza. Należy więc wziąć pod uwagę ten aspekt przy konstrukcji odpowiedniego modelu. Ostatnim, bardzo praktycznym wnioskiem, który można wyciągnąć na podstawie wyników badania, jest to, że niezależnie od metody doboru zmiennych predykcyjnych, wielkości próby badawczej oraz horyzontu prognozy, metoda lasów losowych charakteryzuje się najwyższą zdolnością predykcyjną spośród wszystkich porównywanych metod. Należy pamiętać, że wnioski te zostały wyciągnięte na podstawie próby badawczej 180 spółek akcyjnych, a co za tym idzie – nie powinno się im przypisywać cech uniwersalności.

Zdaniem autora w przyszłych badaniach należy rozszerzyć przeprowadzoną analizę empiryczną. Ciekawym obszarem, ostatnio bardzo popularnym, są hybrydowe modele klasyfikacyjne. Wykorzystane zostały one w badaniach m.in. przez C.L. Chuanga³⁰. Warto sprawdzić, czy ich złożony charakter znajduje odzwierciedlenie w ich wyższej zdolności predykcyjnej.

dr Sergiusz Herman
Uniwersytet Ekonomiczny w Poznaniu
sergiusz.herman@ue.poznan.pl

A COMPARATIVE ANALYSIS OF THE PREDICTABILITY OF SELECTED METHODS FOR PREDICTING BUSINESS FAILURE

Summary

Business failure is a feature of any developed market economy. This phenomenon entails high costs, both economic and social. For this reason, attempts have been made continuously since the beginning of the twentieth century to predict failures of businesses. The interest in this issue is reflected in the application of increasingly advanced statistical methods. The

³⁰ C.L. Chuang, *Application of hybrid case-based reasoning for enhanced performance in bankruptcy prediction*, „Information Sciences” 236, 2013, s. 174–185.

aim of the paper is to compare the predictive capacity of nine methods used in the literature to predict the bankruptcy of enterprises. The empirical research was conducted on the basis of the financial data of 180 Polish public limited companies. Its results made it possible to state that the accuracy of classification of particular methods (and thus their rating) depends on the size of the research sample and on the length of the forecast period. It was also found that the rating of the tested methods does not depend on the chosen method of selection of predictive variables.