

# Jak skuteczne jest sprawozdanie finansowe w przewidywaniu bankructwa firmy?

Maciej Odziemczyk

Wydział Nauk Ekonomicznych, Uniwersytet Warszawski

Maj 2020

## Abstrakt

*Niniejsza praca dotyczy tematyki modelowania prawdopodobieństwa bankructwa przedsiębiorstw z gałęzi produkcyjnej polskiej gospodarki, na danych z lat 2007 - 13. Celem pracy było zbudowanie systemu ostrzegania o bankructwie w horyzoncie rocznym, sprawdzenie czy większe przedsiębiorstwa rzadziej upadają oraz czy uwzględnienie w modelu zmiennej będącej wynikiem modelu Z-score poprawia jego jakość (w tym celu wykorzystano procedurę opartą na symulacji bootstrapowej -  $50 * 1000$  losowań). Finalnie najlepszymi modelami okazały się modele probitowe, estymowane za pomocą Metody Największej Wiarygodności. Ostateczne modele charakteryzowały się odpornością na przetrenowanie, oraz wynikami ROC-AUC score na poziomie 83-84% na próbach testowych i treningowych. Udało się stworzyć zadowalający wynikami system, cechujący się specyficzną na poziomie 95% i wrażliwością 45 - 46% przy ogólnej skuteczności na poziomie 91-92%. Finalnie nie udało się obronić tylko hipotezy o przydatności zmiennej Altmana, odpowiedź na to pytanie okazała się zależna od preferencji badacza.*

## 1 Wstęp

Ludzkość zainteresowana jest problematyką opisu przyczyn bankructwa od bardzo dawna, bowiem jest to zjawisko niosące za sobą olbrzymie skutki, zarówno dla otoczenia mikro (jeżeli są to incydenty), jak i makro (kiedy skala jest większa). Nie wydaje się zatem dziwne, że możliwość przewidzenia takiego zdarzenia jest wysoce atrakcyjna dla wielu podmiotów funkcjonujących w tym ekosystemie, wpływ wiedzy o przyszłym bankructwie na decyzje ekonomiczne jest czymś oczywistym. Z uwagi na fakt nieustannej aktualności problemu, podejmowano wiele prób jego

modelowania, bo jak wiadomo, dobry model, pomimo że jest uproszczoną wersją rzeczywistości, to jest jej esencją i potrafi odpowiedzieć na kluczowe pytania. Działalność przedsiębiorstwa może zostać skwantyfikowana za pomocą licznych wskaźników, mających swoje fundamenty w teorii finansów, rachunkowości i ekonomii. Wskaźniki te zazwyczaj są łatwo dostępne dla wybranych obiektów, jeżeli tylko posiadamy dostęp do ich sprawozdania finansowego, niestety (z punktu widzenia analityka) standardy sprawozdawcze to sprawa wysoce zróżnicowana. Ciągła chęć doskonalenia regulacji, co wynika często z bieżących potrzeb, prowadzi do heterogeniczności szablonów, niekiedy nawet na terenie jednego kraju. Ponadto istnieją podmioty, dla których regulacje w tym zakresie są odrębne (inaczej wygląda sprawozdanie finansowe firmy produkcyjnej, a inaczej instytucji finansowej). Mając na uwadze niejednorodność źródeł danych można stwierdzić, że stworzenie jednego modelu, który zawsze będzie się sprawdzał jest niemożliwe.

Modelowanie bankructwa opiera się na założeniu istnienia dodatniej korelacji między słabą sytuacją finansową, a prawdopodobieństwem, że dana firma zbankrutuje. Problem pojawia się już na samym początku, ponieważ definicji "słabej sytuacji finansowej" i "bankructwa" jest wiele. Pierwsze z zagadnień opisywane jest przez teorię; z reguły w kontekście problemów finansowych pojawiają się takie zjawiska jak ujemny wynik finansowy lub jego spadek na przestrzeni czasu oraz nieadekwatna struktura finansowania i majątkowa. Bankructwo, w zależności od potrzeb, przez różne podmioty może być odbierane inaczej, formalnie według polskiego prawa, za bankruta uznaje się upadły podmiot gospodarczy, tj. taki, który traci zdolność do spłaty swoich wymagalnych zobowiązań pieniężnych, jeżeli opóźnienia przekraczają 3 miesiące lub jeżeli jego zobowiązania pieniężne przekraczają majątek przez okres 24 m-cy. Bankructwo jest zgłaszane przez przedsiębiorcę w ciągu 30 dni, odkąd powstała do tego podstawa prawna [1]. Już samo istnienie terminów - 3 miesiące, 30 dni, 24 miesiące, w konfrontacji z terminem dostępności sprawozdań finansowych (z reguły na koniec roku) prowadzi do powstania kolejnych problemów w modelowaniu. Wadą sprawozdania w tym kontekście jest jego specyfika, jest to coś w rodzaju "zdjęcia" danego podmiotu na koniec roku, jednocześnie jest to najlepsza informacja jaką dysponujemy, ponieważ podlega ona wielu regulacjom i jest audytowana. Dodatkowo mnogość pozycji w dokumentach typu Bilans, Rachunek Zysków i Strat czy Rachunek Przepływów Pieniężnych prowadzi do powstawania jeszcze większej liczby miar. Ze względu na dynamikę zjawiska i jego zróżnicowanie zarówno między-, jak i wewnątrzobiektywne (obiekt jako Państwo), nie zostały nigdy określone ściśle ramy dobrego modelu. Analiza powinna się zatem zacząć od określenia gospodarki, dla której będzie przeprowadzana, a nawet sektora (produkcyjny, finansowy).

#### *Główna hipoteza badawcza*

Celem niniejszej pracy jest sprawdzenie na tle licznych wcześniejszych studiów przeprowadzanych na Świecie, czy dla gospodarki polskiej możliwe jest zbudowanie systemu ostrzegania o bankructwie. Za sprawny system rozumie się możliwe proste do zaaplikowania narzędzie, za pomocą którego

jesteśmy w stanie z wysoką skutecznością (95%) stwierdzić, że obiekt nie zbankrutuje w ciągu roku oraz wykryć przynajmniej 1/3 przedsiębiorstw, które zbankrutowały w tym horyzoncie. Formalnie problem można zapisać w sposób następujący:

$$\begin{cases} P(\hat{p}_i > p^* | y_i = 1) \geq 1/3 \\ P(\hat{p}_i < p^* | y_i = 0) \geq 0.95 \end{cases}$$

gdzie  $\hat{p}$  oznacza oszacowane przez model prawdopodobieństwo przyjęcia przez zmienną zależną wartości 1,  $y_i$  rzeczywistą wartość zmiennej zależnej (1 w przypadku bankructwa, 0 w p.p.), a  $p^*$  oznacza optymalny punkt odcięcia, czyli zdefiniowany poniżej punkt, który przy okazji minimalizuje liczbę błędnych klasyfikacji.

$$\hat{y}_i = \begin{cases} 1 & \text{dla } \hat{p} > p^* \\ 0 & \text{w p.p.} \end{cases}$$

gdzie  $\hat{y}_i$  jest predykcją wartości zmiennej zależnej

#### *Poboczna hipoteza badawcza I*

Duże przedsiębiorstwa mają mniejsze prawdopodobieństwo upadku. Intuicja podpowiada prawdziwość założenia. Ze względu na fakt wywołania większych skutków w przypadku bankructwa szczególnie dużych przedsiębiorstw, otoczenie może prowadzić działania mające zapobiec wystąpieniu tego zdarzenia. Konsekwencje objawiają się chociażby w utracie znacznej liczby miejsc pracy lub produktu, który z jakiegoś powodu jest wartościowy dla społeczeństwa, przy założeniu, że produkt o wyższej zbiorczej użyteczności społeczeństwa jest produktem, na który jest zwiększony popyt. Formalnie hipotezę można zapisać jako

$$\frac{\partial \mathbb{E}[y|x]}{\partial \text{Rozmiar}} < 0$$

#### *Poboczna hipoteza badawcza II*

Jeden z pierwszych i najbardziej popularnych modeli w literaturze przedmiotu to model dyskryminacyjny Z-score Altmana (1968) [2]. Bi-Huei Tsai (2013) wykazała przydatność zmiennej równej punktacji Z-score w swoim modelu wielomianowym logitowym dla chińskiej gospodarki [10]. Owa przydatność stanowi II hipotezę poboczną niniejszej pracy, z tą różnicą że weryfikowana będzie na dysponowanej próbie pochodzącej z gałęzi produkcyjnej polskiej gospodarki. Przydatność zdefiniowana jest jako poprawa pewności predykcyjnej. Formalnie, hipotezę można zapisać:

$$EMC_{standard} > EMC_{Altman}$$

gdzie EMC oznacza ekonomiczny koszt błędu klasyfikacji (*economic misclassification cost*) i można go zapisać jako

$$EMC = y_i * \theta_1 * P(\hat{p}_i < p^* | y_i = 1) + (1 - y_i) * \theta_0 * P(\hat{p}_i > p^* | y_i = 0)$$

gdzie  $\theta_1$  oznacza ekonomiczny koszt błędu I rodzaju tj. klasyfikacji "1" jako "0", a  $\theta_0$  jest ekonomicznym kosztem błędu II rodzaju, tj. klasyfikacji "0" jako "1",  $y_i$  w tym przypadku oznacza zaobserwowaną proporcję sukcesów w próbie, model Standard to model nieuwzględniający zmiennej Altmana w przeciwieństwie do modelu Altman.

Po niniejszym wstępie, w sekcji drugiej dokonany został przegląd literatury w kontekście hipotez badawczych. Opisane zostały po krótko klasyczne artykuły jak i nowsze badania, szczególny nacisk położony został na parametry wrażliwości i specyficzności modeli oraz wykorzystane zmienne, dane i sposób ich obróbki. Sekcja trzecia zawiera opis danych, ich obróbkę oraz przeprowadzoną analizę. W sekcji czwartej przedstawione zostały metody estymacji i metodologia weryfikacji hipotez badawczych. Na sekcję piątą składają się wyniki i wnioski wynikające z przeprowadzonych estymacji i testów. Ostatnia sekcja, szósta, to wyszczególnione główne wnioski z niniejszej pracy oraz zakończenie. W załączniku znajduje się kod źródłowy analizy dokonanej w języku Python w postaci notatnika Jupyter Notebook oraz surowego kodu.

## 2 Przegląd literatury

Od początku historii modelowania bankructwa jako cel główny stawiana była predykcja, opis zjawiska w sensie jaki znamy z analizy ekonometrycznej zazwyczaj był drugoplanowy. Podejście wydaje się być uzasadnione, bowiem informacja o tym czy przedsiębiorstwo zbankrutuje czy nie ma kluczowe znaczenie dla wielu aktorów gospodarki. Pierwsze modele predykcyjne opierane były tylko na sprawozdaniach finansowych i nie wykorzystywały aparatu ekonometrycznego. W latach 70 najpopularniejsza była wieloczynnikowa analiza dyskryminacyjna (*MDA - Multivariate Discriminant Analysis*), a najbardziej rozpoznawalnym modelem tego typu jest Model Z-score Altmana z 1968 roku.

Altman (1968) do swojej analizy wykorzystywał tylko wskaźniki finansowe z obszaru zyskowności, płynności, wypłacalności, efektywności i dźwigni [2]. Spośród 22 zmiennych, na podstawie procedury opartej o potencjał predykcyjny (jednoczynnikowa analiza), korelację, istotność (różnice średnich poszczególnych wskaźników w separowanych grupach) i iteracyjną weryfikację modeli, wybrane zostało 5 wskaźników. Końcowa postać modelu to:

$$Z = 0.12X_1 + 0.14X_2 + 0.033X_5 + 0.006X_4 + 0.999X_3$$

gdzie  $X_1$  = Kapitał obrotowy/Aktywa,  $X_2$  = Zyski zatrzymane/Aktywa,  $X_3$  = EBIT/Aktywa,  $X_4$  = Wartość rynkowa/Wartość księgowa długu,  $X_5$  = Sprzedaż/Aktywa. Warto odnotować, że model szacowany był na próbie konstruowanej na zasadzie "dobierania w pary", która zawierała w sobie 33 bankrutów i 33 nie bankrutów pochodzących ze Stanów Zjednoczonych z lat 1946 -

1965. Kryterium doboru był rozmiar, mierzony sumą aktywów (granice 1 mln\$ - 25 mln\$) i branża produkcyjna. Wyniki Altmana były zdumiewające, na próbie "inicjacyjnej", czyli takiej, która została wykorzystana do estymacji parametrów (teraz nazwalibyśmy ją próbą treingową) uzyskał on skuteczność na poziomie 95% w przypadku jednorocznych predykcji i 72% w horyzoncie dwuletnim. Fantastyczne rezultaty udało się również uzyskać na próbie "testowej" składającej się z 25 przedsiębiorstw o podobnych rozmiarach. Na 25 obserwacji model pomylił się tylko raz, co daje 96% skuteczność. Model został przetestowany na jeszcze jednej próbie 66 firm i uzyskał tam 79% skuteczność. Próbowano też predykcji na odleglejszych horyzontach czasowych, jednak rezultaty nie były już tak spektakularne. Generalnie model działał na zasadzie "im większy Z-score tym mniejsze prawdopodobieństwo bankructwa", a za wartość krytyczną uznano 2.675. Swym artykułem Altman obronił tezę o skuteczności wskaźników finansowych w modelowaniu bankructwa, jednakże nie sposób nie zwrócić uwagi na możliwe bardzo duże przetrenowanie modelu.

Ohlson (1980) jako pierwszy zaproponował model logistyczny do celu predykcji bankructwa przedsiębiorstw; jest on zatem prekursorem czysto ekonometrycznego podejścia do zagadnienia [8]. W swoim artykule szacował on model na znacznie większej próbie niż poprzednicy. Badania Ohlsona opierały się na 2058 niebankrutach i 105 bankrutach, a dane pochodziły ze sprawozdań finansowych. Zasotowane podejście ma kilka istotnych zalet np. rozkład bankrutów w próbie jest bliższy rozkładowi w populacji, podczas gdy w badaniu Altmana to było 50/50, ponadto uniknięto procedury "dobierania w pary" dzięki czemu Metoda Największej Wiarygodności mogła być w ogóle zastosowana, bo jak wiadomo jednym z jej założeń jest losowość próby (niezależność obserwacji). W swoim modelu Ohlson wykorzystał zmienne takie jak rozmiar spółki, mierzony logarytmem stosunku Aktywów do indeksu cen (-), stosunkiem Zobowiązań i Aktywów (+), Kapitału obrotowego i Aktywów (-), Wskaźnika płynności bieżącej (+), ROA (-), EBIT/Aktywa (-), Przyrost zysku (-) i Zmienne binarne, jedną mówiącą o tym czy w przeciągu dwóch ostatnich lat wystąpiła Strata (+), druga, czy Zobowiązania przekroczyły Aktywa (zaakcentowanie wykorzystanej zmiennej ciągłej) (+). Dodatkowo za punkt odcięcia maksymalizujący skuteczność uznany został punkt 0.038 (Specyficzność 0.826 i wrażliwość 0.876). Swoją Artykuł Ohlson podsumował krótko: wskaźniki są odporne na metody estymacji i jeżeli chce się uzyskać lepsze wyniki należy poszerzyć zbiór zmiennych objaśniających.

Ostatnim z klasyków literatury przedmiotu jest Zmijewski (1984), on z kolei jako pierwszy zaproponował model probitowy [13]. W swoim artykule Aurotr podniósł problem obciążenia estymatorów poprzez sztuczne bilansowanie próby (wyrównywanie "1" i "0", a przynajmniej niezachowywanie proporcji zbliżonych do populacji) oraz nielosowy jej dobór. Kolejnym powodem dla którego "Methodological Issues Related to the Estimation of Financial Distress Prediction Models" jest uznawany za wpływowy artykuł jest fakt zwrócenia uwagi na możliwość wykorzystania informacji o braku informacji. Takie podejście wydaje się uzasadnione w modelowaniu zjawiska

tak "śliskiego" i niepożądanego jak bankructwo, przedsiębiorstwa mogą bowiem próbować zatajać pewne fakty. Próba, którą dysponował Zmijewski pochodziła z lat 1972 - 78 i zawierała spółki z nowojorskiej giełdy (2080-2241 rocznie). Model był prosty, zawierał bowiem tylko 3 zmienne objaśniające ROA (-), Zadłużenie/Aktywa (+) i Wskaźnik płynności bieżącej (-), z czego ten ostatni, podobnie jak u Ohlsona okazał się nieistotny statystycznie, dodatkowo oszacowany został przeciwny znak. Ważone MNW zaproponowane przez Zmijewskiego polegało na przemnożeniu komponentów funkcji wiarygodności przez stosunek "1" i "0" w populacji (podejście wskazane jako właściwe w przypadku nieodpowiednich proporcji w próbie w stosunku do populacji).

Darush Yazdanfar (2011 jako bardziej współczesna autorka, korzystająca z dorobku poprzedników zwróciła uwagę na brak teoretycznych ram wskazujących na najlepszy sposób badania zjawiska) [12]. Jako cel artykułu postawione zostało znalezienie grupy wskaźników finansowych, które w zadowalającym stopniu potrafią różnicować przedsiębiorstwa zagrożone bankructwem od tych w dobrej kondycji dla szwedzkiej gospodarki. Dane służące do estymacji pochodziły z biznesowej szwedzkiej bazy danych i zawierały 75553 obiekty w dobrej kondycji i 3064 bankrutów z lat 2004 - 06. Bardzo ważna jest tutaj kwestia obróbki danych, gdyż reżim narzucony przez Autorkę był bardzo surowy i doprowadził do znacznego okrojenia próby. Yazdanfar zdecydowała się na usunięcie wszystkich obserwacji zawierających braki w danych oraz wszystkie te, które różniły się o +/- 3 rozstępy międzykwartylowe od kwartyli 1 i 3. Zaprezentowane podejście skutkowało zmniejszeniem próby do 3084 aktywnych nie bankrutów i 1412 bankrutów. Do wyestymowanego metodą krocząca modelu regresji logistycznej postawiony został szereg prostych hipotez, każda o znaku oszacowania dla wskaźnika z odpowiedniej grupy (płynności, dźwigni, zadłużenia itp.). Finalne oszacowania dla jednorocznych predykcji (bowiem zostały zbudowane również modele dla dwu- i trzyletniego horyzontu) zgodne były z oczekiwaniami i prezentowały się następująco: Zobowiązania krótkoterminowe/Aktywa (+), Zobowiązania/Aktywa (-), Względny przyrost aktywów (-), Rozmiar, mierzony logarytmem naturalnym sprzedaży (-), Koszty finansowe/Zobowiązania (+), ROE (-). Ze względu na wykorzystanie metody selekcji kroczącej, wszystkie parametry były statystycznie istotne, co więcej nie udało się znaleźć innych wskaźników poprawiających wyniki dla tej próby. Wartym odnotowania jest też fakt dodatniego znaku przy stałej, co jest niecodziennym zjawiskiem w badaniach nad tematem bankructwa, z reguły jest to -. Skuteczność predykcji opiewała natomaist na 96.7% (wrażliwość 0.933 i specyficzność 0.983). Zaobserwowany został również spadek skuteczności, wraz z wydłużaniem horyzontu.

Bardzo ciekawe badanie zostało zaprezentowane w 2013 roku przez Bi-Huei Tsai - polegało ono na weryfikacji hipotez o równoważności modeli standardowych, opartych tylko na wskaźnikach finansowych, modeli zawierających dodatkowo zmienne rynkowe oraz modeli zawierających w sobie oryginalną punktację Z-score Altmana z 1968 roku, dodatkowo podjęty został temat doboru optymalnego punktu odcięcia [10]. Próba zawierała w sobie wszystkie przedsiębiorstwa notowane

na tajwańskiej giełdzie, z wyłączeniem instytucji finansowych (inne zasady rachunkowości) oraz firm z brakami kluczowych atrybutów. Dane pochodziły z lat 2003 - 09 i zawierały w sobie 40 bankrutów, 147 firm klasyfikowanych jako posiadających problemy finansowe oraz 4143 dobrze prosperujące jednostki. Autorka zwróciła uwagę na złożoność problemu i jego niekoniecznie dychotomiczną specyfikę, dlatego próba zawierała trzy stany, a estymowany był wielomianowy model logitowy. Kluczowym elementem badania nie była jednak bezpośrednio skuteczność predykcji, a sprawdzenie czy dodanie zmiennej Altmana wpływa na poprawę pewności predykcji. Owa miara pewności oparta była na doborze optymalnego punktu odcięcia zaprezentowanego przez Granger i Peserana w 2000 roku i obliczeniu kosztu błędu klasyfikacji dla bootstrapowej symulacji (1000 losowań dla pojedynczego punktu odcięcia z przedziału 0.01 - 0.3). Jako optymalny punkt odcięcia do wykrywania firm z "przejściowymi problemami" wybrany został punkt 0.02, natomiast dla klasyfikacji bankrutów była to wartość 0.17. Wyniki były interesujące, bowiem model zawierający zmienną Altmana okazał się zwycięzcą konkursu, notując najniższy koszt błędu klasyfikacji. Ponadto podkreślona została również wyższa skuteczność w wykrywaniu bankrutów niż firm z problemami finansowymi, co może sugerować że predykcje są bardziej wiarygodne dla problemów binarnych.

Bezpośrednio do artykułów Altmana i Ohlsona odnieśli się Julian Ni, Wikil Kwak, Xiaoyan Cheng i Guan Gong w swojej pracy (2014), dotyczącej predykcji bankructwa przedsiębiorstw gospodarki chińskiej [6]. Badacze dysponowali próbą, a w zasadzie populacją przedsiębiorstw z lat 1999 - 2007. Z finalnej próby wyłączone zostały obserwacje z lat 2001 - 04, ze względu na niemożność wydzielenia grupy bankrutów i nie bankrutów. Ponadto usunięte zostały obserwacje z brakami w danych, oraz te które uznane zostały przez Autorów za błędne (podany przykład ujemnych Aktywów). Dodatkowo zdecydowano się również na selekcję obiektów do próby na zasadzie "dobierania w pary" w stosunku 2:1 i tak, finalnie otrzymano próbę składającą się z 521 bankrutów i 950 (podane w artykule) dobrze prosperujących przedsiębiorstw z lat 1999 - 2007 (z wyłączeniem lat 01 - 04). Autorzy zdecydowali się na estymację kilku modeli logistycznych - pierwszy opierał się na zmiennych zaproponowanych przez Altmana (1968), drugi na zmiennych zaproponowanych przez Ohlsona (1980) i trzeci, będący połączeniem dwóch poprzednich. Ponadto, w każdym z modeli uwzględniona została struktura własnościowa jako zmienne binarne. Wyniki nie były ściśle spójne z pierwowzorami, bowiem w modelu Altmanowskim stosunek Wartości rynkowej i Wartości księgowej długu miał znak dodatni, podobnie jak Sprzedaż/Aktywa (ta zmienna podobnie jak u Altmana, została wykazana jako nieistotna statystycznie), a w modelu Ohlsonowskim, różniły się istotności zmiennych, znaki natomiast były tożsame. Model łączny zachował znaki takie jak modele szczególne. Warto odnotować fakt nieistotności Rozmiaru mierzonego stosunkiem Aktywów do PKB w danym roku. Wyniki predykcyjne były następujące: 72.26% skuteczności, przy 33.75% wrażliwości i 93.37% specyficzności dla modelu Altmanowskiego, 73.89% skuteczności ogólnej,

36.08% wrażliwości i 94.63% specyficzności dla modelu Ohlsonowskiego oraz 74.64% skuteczności przy 39.92% wrażliwości i 93.68% specyficzności dla modelu ogólnego.

Jacek Welc w swoim artykule z 2017 roku postanowił sprawdzić, czy możliwa jest predykcja oparta o model logitowy zawierający tylko jedną zmienną poza stałą [11]. Poza sprawdzeniem skuteczności, artykuł miał odpowiedzieć na pytanie czy wskaźnik oparty o rachunek przepływów pieniężnych jest lepszym predyktorem. Przesłanką była potencjalna odporność rachunku przepływów pieniężnych na manipulacje sprawozdaniem finansowym, w przypadkach przedsiębiorstw zagrożonych bankructwem. Testowane wskaźniki to EBITDA/Zobowiązania jako reprezentant klasycznego podejścia oraz Operacyjne przepływy pieniężne/Zobowiązania jako reprezentant RPP. Próba pochodziła z polskiej gospodarki z lat 2009 - 16 i zawierała 92 losowo wybrane spółki, które nie zbankrutowały i 92 losowo wybrane spółki, które zbankrutowały. Skuteczność modeli opiewała na 66% - 76%, ponadto nie stwierdzono wyższości Rachunku Przepływów Pieniężnych w analizie.

Rok po artykule Welca, dwójka badaczy, Hamid Waqas i Rohani Md-Rus postanowiła powtórnie spojrzeć na możliwości Rachunku Przepływów Pieniężnych, tym razem skupiając się jednak na zbudowaniu kompleksowego modelu dla rynku pakistańskiego [5]. Celem artykułu było również sprawdzenie generalizowania się modelu Ohlsona na gospodarki rozwijające się za pomocą porównania mocy predykcyjnej tego modelu z autorskim. Model autorski uwzględniał zmienne takie jak ROA (+), Zyski zatrzymane/Aktywa (-), EBIT/Aktywa (-), Current ratio (+), Kapitał obrotowy/Aktywa (-), Zobowiązania/Aktywa (+) i jako reprezentanci Rachunku Przepływów Pieniężnych - stosunek Operacyjnych przepływów pieniężnych do Zobowiązań (+) oraz Sprzedaży (+). Autorzy podkreślili wypełnienie luki w literaturze poprzez skorzystanie ze wskaźników opartych na RPP w towarzystwie tych powszechnie używanych, estymowanie modelu dla rynku rozwijającego się oraz w kontekście literatury krajowej, poprzez stosunkowo dużą próbę treningową nieograniczoną co do branży. Wspomniana próba liczyła 290 obserwacji, z czego liczba bankructów wyniosła 45. Dane pochodziły z lat 2006 - 16 i dotyczyły spółek notowanych na tamtejszej giełdzie. Model Ohlsona zanotował 61% skuteczności przy wrażliwości 85% i specyficzności 56% - co jest dość niespotykanym wynikiem, zazwyczaj modele są bardziej specyficznie niż wrażliwe. Model autorski natomiast to już skuteczność na poziomie 91%, przy 56.4% wrażliwości i 98.1% specyficzności. Wnioski z badania zatem są takie zbliżone do tych, które wyciągnął Ohlson, chcąc otrzymać lepsze wyniki, należy poszerzyć zbiór zmiennych objaśniających, dodatkowo pokazana została konieczność estymacji modelu na danych pochodzących z rynku, który chcemy przewidywać.

Gigantyczną bazą danych dysponowali natomiast Charalambakis i Garrett (2019) [4]. Ich próba zawierała 30886 obserwacji, z czego bankructów było 1770, a zbiór zmiennych objaśniających, ze względu na specyfikę bazy danych zawierał tylko wskaźniki finansowe (brak zmiennych rynkowych). Najbardziej kompleksowy model poza wskaźnikami finansowymi zawierał również



zmienne binarne mówiące o tym czy firma jest eksporterem, czy wypłaca dywidendy oraz czy dane pochodzą z okresu kryzysu gospodarczego w Grecji. Dodatkowym urozmaicheniem było zaprezentowanie metryki ROC-AUC score. Najlepszy model (logitowy), uzyskał wynik ROC-AUC na poziomie 0.8195, co na tle estymatora losowego, którego wynik to z definicji zawsze 0.5 oznacza, że model dość dobrze nadaje się do predykcji. Oszacowany w pracy model bardzo dobrze wykrywał bankrutów - dla 87.6% z nich dopasowane prawdopodobieństwo mieściło się w decylach 1-5 na próbie treningowej oraz 88.34% na próbie testowej, co oznacza, że ze standardowym progiem odcięcia 0.5 model charakteryzował się wrażliwością na poziomie 87-88%, niestety specyficzność nie została opisana w artykule. Wydłużając horyzont czasowy, uzyskano gorsze, ale dalej bardzo dobre wyniki, 2 lata to wrażliwość 86.25% na próbie testowej, 3 lata to 83.45%. Co ciekawe, podobnie do przypadku rocznego horyzontu, model lepiej spisywał się na próbach testowych niż treningowych (wrażliwość, 2 lata - 83.52%, 3 lata - 80.9%).

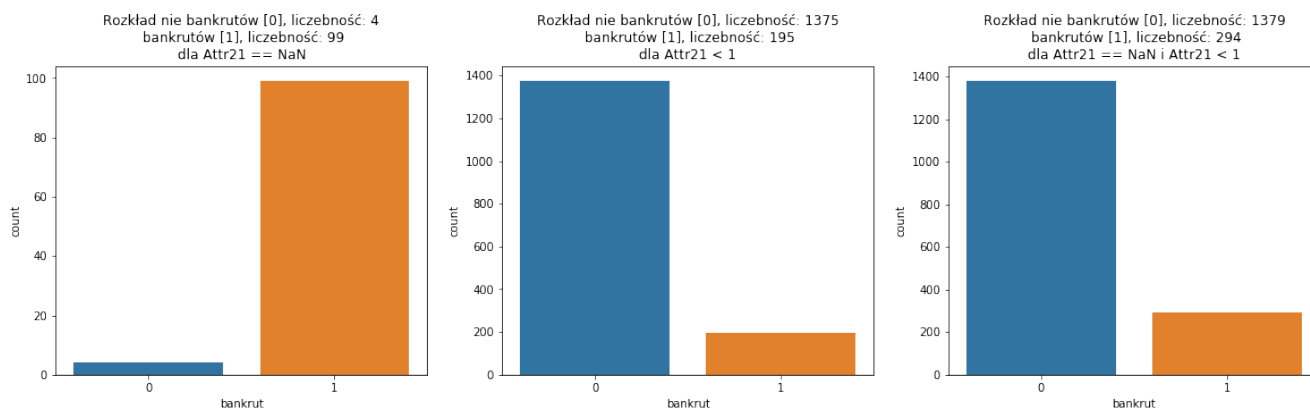
W 2019 roku, Christopoulos, Dokasb, Kalantonisc oraz Koukkoud na próbie spółek, z sektora chemicznego i metalurgicznego notowanych na nowojorskiej giełdzie, wyestymowali model hazardowy oparty o dynamiczną specyfikację logitową [3]. Celem badania było sprawdzenie hipotezy o tym, że wskaźniki zyskowności i płynności są kluczowymi kryteriami rozróżniającymi spółki dobrze prosperujące od tych z problemami (*financial distress*, kryteria zostały określone w toku rozmów z ekspertami oraz badaczami uniwersyteckimi), oraz sprawdzenie, czy status *financial distress* zmienił się po czasie wielkiego kryzysu lat 2007 - 09. Autorzy zdecydowali się przyjąć konwencję oszacowania 3 modeli, pierwszy opierał się tylko na wskaźnikach płynności, drugi tylko na wskaźnikach zyskowności, a trzeci stanowił kombinację dwóch poprzednich. Przeprowadzone badanie znacznie różniło się metodologicznie od poprzednich, zaaplikowane zostały takie techniki jak bootstrap oraz walidacja krzyżowa, w toku których dokonano doboru prób treningowych (110, 53 "1" i 57 "0") i testowych (28) oraz estymacji po uprzednim podziale próby na dwie podpróby (osobno przemysł metalurgiczny - głównie wydobywanie złota i farmaceutyczny) . Ciekawy był również sposób obróbki danych - wykorzystano zakres 1 - 99 centyli ze zbiorczych rozkładów. Ostatni model (zawierający zmienne mówiące zarówno o płynności, jak i o zyskowności) cechował się skutecznością na poziomie 82.51%, z czego błąd I rodzaju (klasyfikacja "1" jako "0") wyniosła 17.58%, a błąd II rodzaju (klasyfikacja "0" jako "1") wyniosła 1.11%, przy wyniku ROC-AUC na poziomie 73.77%. Wykorzystanie procedury bootstrapowej pozwoliło na poprawę wyników skuteczności o 1-2%. Najlepszy okazał się model oparty na wskaźnikach zyskowności, zawierał on dwie istotne zmienne - Marżę zysku brutto, oraz Koszty operacyjne/Przychody. Pomimo, że badanie nie odnosiło się stricte do predykcji bankructwa, a *financial distress*, warto zwrócić uwagę na unikatowość niektórych z wykorzystanych zmiennych objaśniających, które spisały się bardzo dobrze (np. Marża zysku brutto - wykorzystana w modelu będącym przedmiotem tej pracy).

### 3 Dane i ich analiza

Dysponowana próba to gotowy zbiór danych przygotowany przez dr Sebastiana Tomczaka z Katedry Badań Operacyjnych i Inteligencji Biznesowej Wydziału Informatyki i Zarządzania Politechniki Wrocławskiej, na bazie którego powstał również artykuł "Ensemble Boosted Trees with Synthetic Features Generation in Application to Bankruptcy Prediction" autorstwa Macieja Zięby, Sebastiana Tomczaka i Jakuba Tomczaka [7]. Badane obiekty to przedsiębiorstwa z sektora produkcyjnego należące do polskiej gospodarki z lat 2007 - 13 w przypadku bankrutów i 2000 - 12 w przypadku nie bankrutów. Dane pochodzą z bazy serwisu Emerging Markets Information Service (EMIS) zawierającej w sobie informacje na temat rynków wschodzących, w tym Polski (finansowe, polityczne, makroekonomiczne i newsy lokalne). Zmienne objaśniające - oryginalnie 64 wskaźniki finansowe zostały obliczone na podstawie sprawozdań finansowych, a zmienna objaśniana to etykieta - "1" w przypadku zadeklarowania bankructwa, "0" w przeciwnym razie. Oryginalnie zbiór podzielony został na 5 przypadków klasyfikacji, odpowiadających problemom w horyzoncie odpowiednio 5-cio, 4-ro, 3-y, 2-u letnim i 1-no rocznym. Do badania prezentowanego w niniejszej pracy wykorzystano podzbiór o horyzoncie 1-no rocznym. Próba ta zawierała w sobie 5910 obserwacji, z czego w 410 przypadkach przypisana została etykieta bankruta, co oznacza silnie niezbalansowaną, ale wiarygodną w stosunku do populacji pod względem proporcji próbę. Oczywiście surowe dane nie najlepiej nadają się do zastosowania procedury ekonomertycznej, ponadto nie zawierały wszystkich zmiennych, które postanowiono wykorzystać w badaniu, dlatego też dużą część pracy nad modelem stanowiła analiza i obróbka danych.

Braki danych w większości dotyczyły zmiennej mówiącej o względnym przyroście Sprzedaży (101 ze 123 w ogóle), dlatego też zdecydowano się utworzyć z niej zmienną binarną, przyjmującą wartość "1" w przypadku wystąpienia spadku sprzedaży w stosunku do poprzedniego roku lub braku danych oraz "0" w przeciwnym wypadku. Pozostałe 22 obserwacje, co do których nie udało się odnaleźć prawidłowości występowania braków danych zostały usunięte. Powodem takiego kształtu zmiennej binarnej był wynik testu, w którym okazało się, że w przypadku utworzenia zmiennej trzypoziomowej, jeden z poziomów (Brak danych) we wszystkich obserwacjach wiązał się z bankrutami. Poniższy wykres obrazuje zasadność podejścia.

Rysunek 1: Wykresy rozkładu bankrutów i nie bankrutów w zależności od manipulacji zmienną Przyrost sprzedaży



Źródło: Opracowanie własne. Attr21 to zmienna ciągła opisująca przyrost sprzedaży, NaN oznacza brak danych.

Następnie utworzono zmienne:

- dźwignia finansowa jako iloczyn stosunku Przychodów na działalności operacyjnej do Sprzedaży (Attr42 w bazie) i odwrotności stosunku Zysku brutto do Sprzedaży (Attr19 w bazie),
- Z-score Altmana, obliczony według oryginalnej formuły (z tą różnicą, że zamiast Wartości rynkowej posłużono się Wartością księgową kapitału własnego (ze względu na dostępność danych),
- zmienną binarną "wyplacalnosc", przyjmującą wartość "1" w przypadku kiedy  $\text{Zobowiązania} > \text{Aktywa}$ , 0 w p.p.

W celu upewnienia się co do spełnienia założenia o losowości próby, oraz chęci sprawdzenia generalizacji wyestymowanego później modelu zastosowano losowy podział dysponowanej próby na próbę treningową oraz testową w proporcji 80/20. Dalsza analiza odbywała się wyłącznie na zbiorze treningowym w celu zachowania warunków generalizacji najbardziej zbliżonych do rzeczywistości (np. na próbie testowej nie rozpatrywano wartości odstających). Zmienne biorące udział w analizie, wraz z opisem (formułami) zostały zaprezentowane w Tabeli 1.

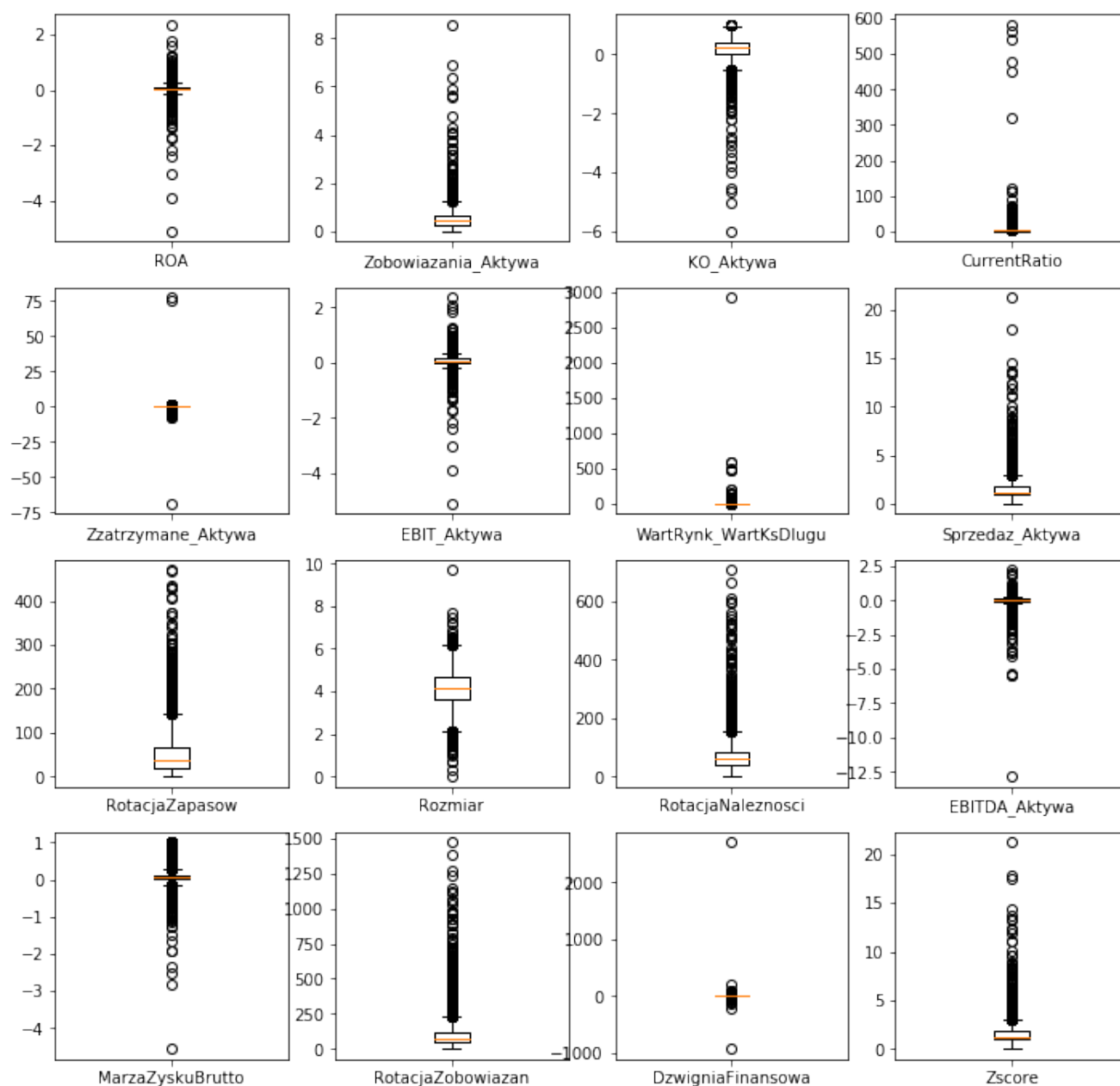
Tablica 1: Zmienne objaśniające

Nazwa zmiennej	opis
ROA	Zysk netto/Aktywa
Zobowiazania_Aktywa	Zobowiązania/Aktywa
KO_Aktywa	(Aktywa obrotowe - Zobowiązania bieżące)/Aktywa
CurrentRatio	Aktywa obrotowe/Zobowiązania bieżące
Zzatrzymane_Aktywa	Zyski zatrzymane/Aktywa
EBIT_Aktywa	Zyski przed odsetkami i podatkami/Aktywa
WartRynkowa_WartKsDlugu	Wartość księgowa kapitału własnego/Wartość księgowa długu
Sprzedaz_Aktywa	Sprzedaż/Aktywa
RotacjaZapasow	(Zapasy * 365)/Sprzedaż
Rozmiar	logarytm naturalny sumy bilansowej
RotacjaNaleznosci	(Należności * 365)/Sprzedaż
EBITDA_Aktywa	(EBIT + Amortyzacja + Deprecjacja)/Aktywa
MarzaZyskuBrutto	(Sprzedaż – Koszty sprzedanych produktów)/Sprzedaż
RotacjaZobowiazan	(Zobowiązania handlowe * 365)/Sprzedaż
DzwigniaFinansowa	Przychody na działalności operacyjnej/Zysk brutto
Zscore	Zmienna Altmana
PrzyrostSprzedazySpadek_BrakDanych	binarna, 1 jeżeli Sprzedaż (t-1) > (t) lub brak danych
wypłacalność	1 jeżeli Zobowiązania/Aktywa > 1

Źródło: Opracowanie własne.

Błędne obserwacje zostały zdiagnozowane na podstawie pojedynczej weryfikacji przypadków w sposób przedstawiony przez Ni i in. (2014) (metoda "ekspercka"). Łącznie na tej podstawie usunięto 36 obserwacji (12 bankrutów, 24 nie bankrutów), w których np. średni czas Rotacji zobowiązań lub należności przekraczał 5 lat. Szczegółowy wykaz obserwacji usuniętych, wraz z uzasadnieniem znajduje się w załączonym notatniku (Jupyter notebook) z kodem źródłowym analizy. Finalnie próba treningowa została okrojona z 4711 obserwacji (308 bankrutów - 6.54%, 4403 nie bankrutów - 93.46%) do 4675 obserwacji (296 bankrutów - 6.33%, 4379 nie bankrutów - 93.77%), próba testowa natomiast zawierała 1177 obserwacji (98 bankrutów - 8.33% i 1079 nie bankrutów - 91.67%) Rozkłady zmiennych ciągłych zostały przedstawione na wykresach pudełkowych poniżej.

Rysunek 2: Wykresy pudełkowe dla zmiennych ciągłych po usunięciu błędnych obserwacji

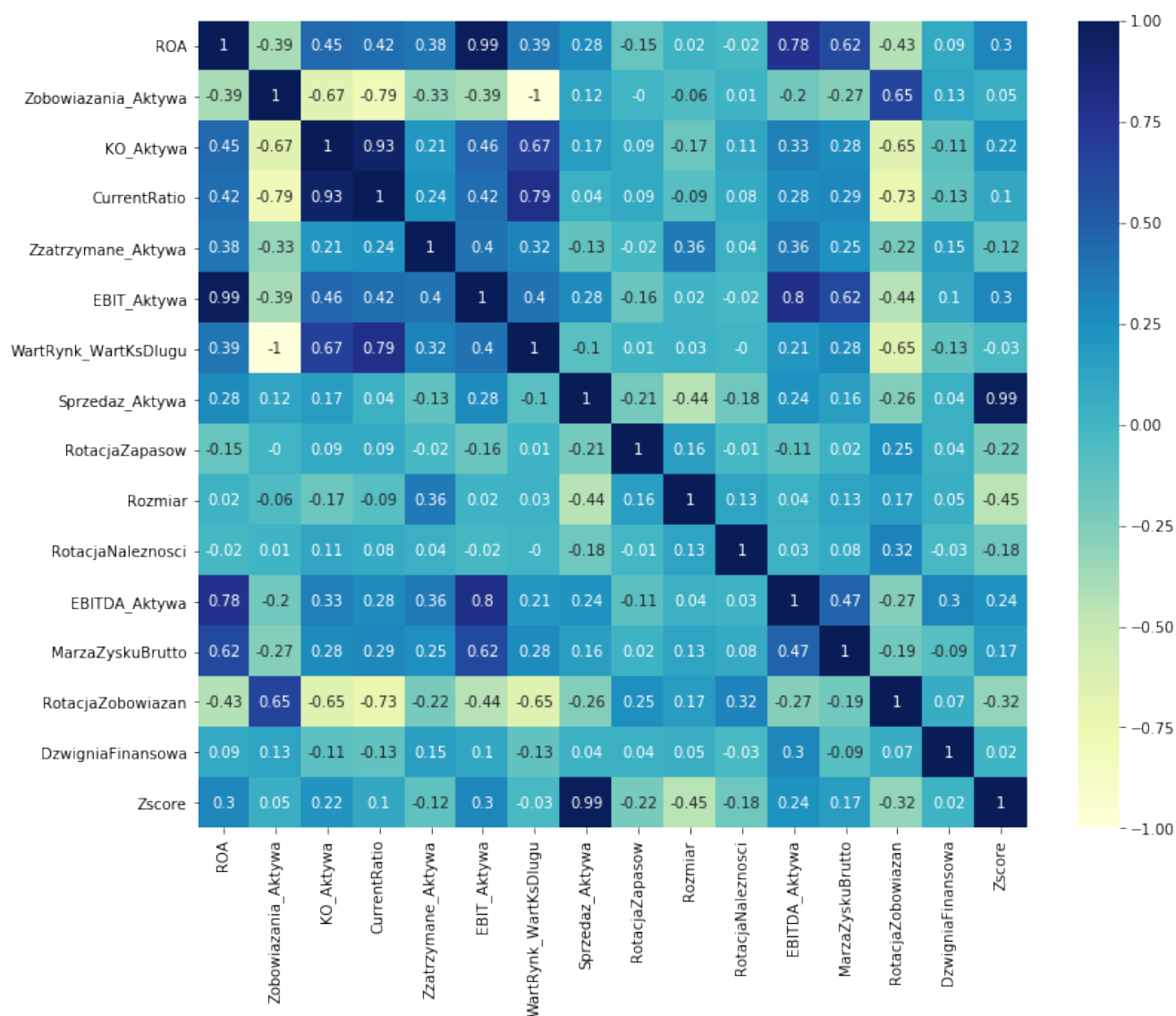


Źródło: Opracowanie własne.

Wykresy pudełkowe bardzo dobrze pokazują, jak bardzo zmienne są zróżnicowane, dzięki czemu można zauważyć, że usunięcie wszystkich tych obserwacji, które różnią się od kwartyli 1 i 3 o trzykrotność rozstępu kwartyłowego (standardowe podejście w analizie danych, wykorzystane przez np. Darush Yazdanfar (2011)) może prowadzić do znacznej utraty informacji, również na dysponowanej próbie treningowej [12]. Jednocześnie wykresy pudełkowe dobrze obrazują typową dla wskaźników finansowych asymetrię ich rozkładów, wzięto ten fakt pod uwagę i przy badaniu zależności między zmiennymi zrezygnowano ze współczynnika korelacji liniowej Pearsona, na rzecz

współczynnika korelacji rang Spearmana. Poniżej heatmapa wspomnianej macierzy korelacji.

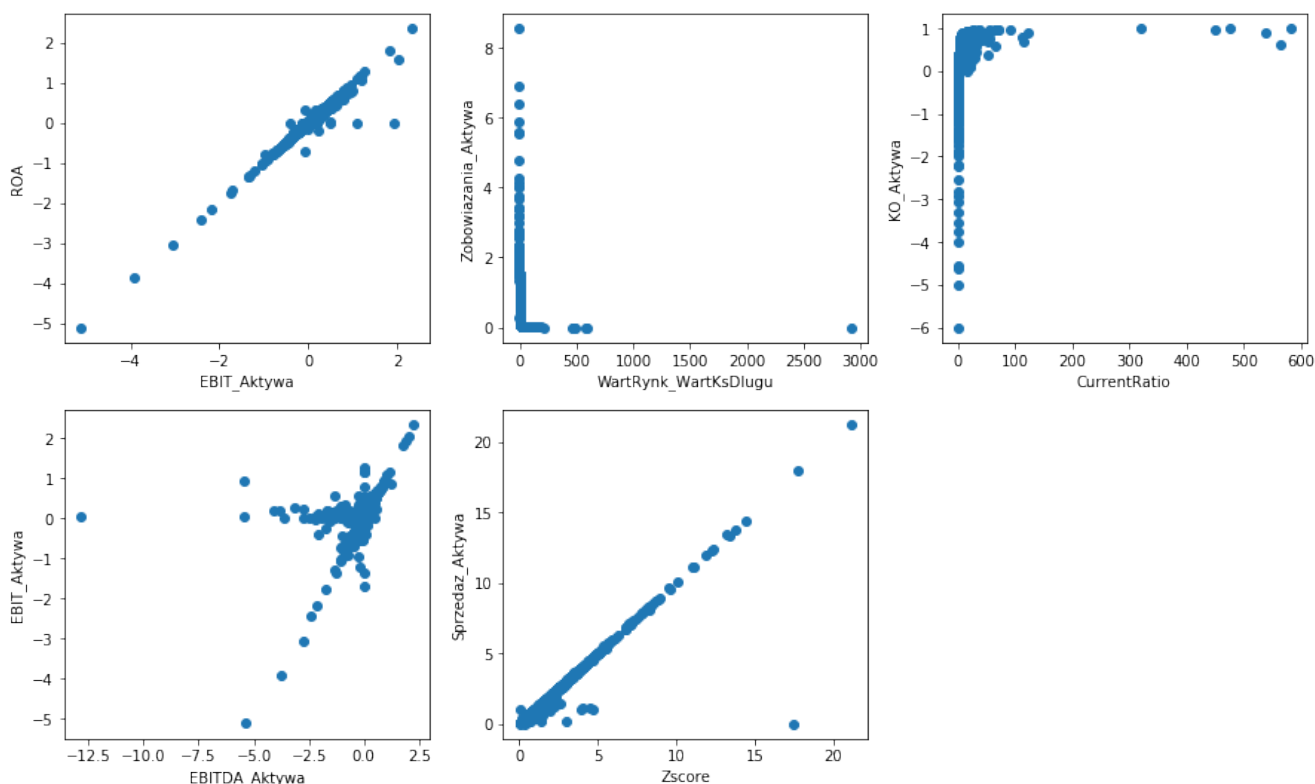
Rysunek 3: Macierz korelacji rang Spearmana (heatmapa)



Źródło: Opracowanie własne.

Z uwagi na fakt wykrywania przez współczynnik korelacji Spearmana również zależności nieliniowych, zdecydowano się na sporządzenie wykresów rozrzutu dla zmiennych skorelowanych ze sobą w stopniu  $> 0.8$ , aby zobaczyć, jakiego rodzaju jest to zależność.

Rysunek 4: Wykresy rozrzutu dla zmiennych skorelowanych  $> 0.8$



Źródło: Opracowanie własne.

W toku analizy uznano zmienną EBITDA\_Aktywa za nieprzydatną, bowiem okazało się, że w 4673 obserwacjach (dla całej bazy danych) zachodziła nierówność  $EBIT > EBITDA$ , a w 3958  $ROA > EBITDA$ , co z definicji EBITDA ( $EBIT + \text{Amortyzacja} + \text{Deprecjacja}$ ) nie może zachodzić na tak dużą skalę (ujemna Amortyzacja i Deprecjacja nie wydaje się być naturalna, zwłaszcza w takim stosunku), dlatego też zdecydowano się usunąć EBITDA\_Aktywa ze zbioru zmiennych objaśniających (dodatkowo silna korelacja ze zmienną EBIT\_Aktywa). Z uwagi na silnie liniowe zależności niektórych zmiennych oraz niepewność co do zachowania modelu w przypadku zobrazowanych nieliniowych zależności, zdecydowano się również sprawdzić występowanie zjawiska współliniowości zmiennych egzogenicznych (*multicollinearity*) za pomocą wskaźników inflacji wariancji (*VIF*). Wyniki przedstawiono w Tabeli 2.  $VIF_1$  to *VIF* policzony dla wszystkich zmiennych za wyjątkiem Zscore (ta zmienna zostanie wyłączona ze zbioru zmiennych objaśniających do modelu ogólnego, z uwagi na treść jednej z hipotez badawczych),  $VIF_2$  to *VIF* policzony po usunięciu zmiennej EBIT\_Aktywa (wyjaśnienie poniżej). Dodatkowo warto nadmienić, że Tabela 2 zawiera wszystkie zmienne, które zostaną wykorzystane do oszacowania modelu ogólnego (interakcje i potęgi, które nie były uwzględnione w Tabeli 1).

Tablica 2: Wskaźniki inflacji wariancji

zmienna	$VIF_1$	$VIF_2$
ROA	26.76	1.65
Zobowiazania_Aktywa	3.25	3.25
KO_Aktywa	3.7	3.69
CurrentRatio	1.17	1.17
Zzatrzymane_Aktywa	1.02	1.02
EBIT_Aktywa	26.92	-
WartRynkowa_WartKsDlugu	1.15	1.15
Sprzedaz_Aktywa	1.60	1.58
RotacjaZapasow	1.33	1.32
PrzyrostSprzedazySpadek_BrakDanych	1.19	1.32
Rozmiar	37.77	37.77
RotacjaNaleznosci	1.41	1.40
MarzaZyskuBrutto	1.25	1.25
RotacjaZobowiazan	2.46	2.45
DzwigniaFinansowa	1.03	1.03
wyplacalnosc	2.30	2.30
Rozmiar <sup>2</sup>	35.88	35.88
wyplacalnosc:PrzyrostSprzedazySpadek_BrakDanych	1.95	1.94

Źródło: Opracowanie własne

W literaturze, VIF na poziomie poniżej 5 - 10 uznaje się za dopuszczalny. Jak można zauważyć w powyższej tabeli ( $VIF_2$ , po usunięciu zmiennej EBIT\_Aktywa), tylko dwie zmienne charakteryzują się VIF wskazującym na współliniowość, są to jednak potęgi tej samej zmiennej - Rozmiar, nie należy się tym zatem posilkować jako przesłanką do usunięcia którejkolwiek z tych zmiennych.



## 4 Metodologia

Do modelowania zjawiska bankructwa w literaturze najczęściej stosuje się modele oparte na Metodzie Największej Wiarygodności (MNW, w podejściu ekonometrycznym; pominięty został tutaj fakt stosowania rozwiązań Sztucznej Inteligencji i Uczenia Maszynowego, oraz przestarzałych już modeli MDA). Podobnie jest w niniejszej pracy, dodatkowo poza modelami opartymi na MNW (logit i probit) podjęta została próba oszacowania Liniowego Modelu Prawdopodobieństwa (LPM) za pomocą Metody Najmniejszych Kwadratów z wykorzystaniem odpornej na heteroskedastyczność macierzy odpornej White'a (LPM charakteryzuje się heteroskedastycznością).

Metoda Największej Wiarygodności polega na maksymalizacji tzw. funkcji wiarygodności (a w praktyce, dla wygody, jej logarytmu - ze względu na monotoniczność przekształcenia punkt maksimum jest ten sam). Funkcja wiarygodności ma postać:

$$L_{\theta}(\mathbf{y}|\mathbf{X}) = \prod_{i=1}^N f_{\theta}(y_i|\mathbf{x}_i)$$

a jej logarytm:

$$l_{\theta}(\mathbf{y}|\mathbf{x}) = \sum_{i=1}^N \ln f_{\theta}(y_i|\mathbf{x}_i)$$

gdzie  $f_{\theta}$  oznacza warunkowe prawdopodobieństwo zaobserwowania wartości zmiennej endogenicznej.

Postać funkcji wiarygodności implikuje zatem konieczność spełnienia założenia o niezależności (losowości) próby.

Dla zmiennych binarnych prawdopodobieństwo zajścia pojedynczego zdarzenia dane jest rozkładem dwumianowym:

$$Pr(y_i|\mathbf{x}_i) = \begin{cases} 1 - p(\mathbf{x}_i) & \text{dla } y_i = 0 \\ p(\mathbf{x}_i) & \text{dla } y_i = 1 \end{cases} = [p(\mathbf{x}_i)]^{y_i} * [1 - p(\mathbf{x}_i)]^{1-y_i}$$

Zatem, logarytm funkcji wiarygodności można zapisać jako:

$$l(\boldsymbol{\theta}) = \sum_{i=1}^N y_i * \ln p(\mathbf{x}_i) + \sum_{i=1}^N (1 - y_i) * \ln[1 - p(\mathbf{x}_i)]$$

W praktyce, modele dla zmiennej binarnej estymowane za pomocą MNW różnią się właśnie funkcją  $p(\cdot)$ , która jest dystrybuantą  $F(\cdot)$  odpowiedniego dla modelu rozkładu:

- Probit  $F(\mathbf{x}_i) = \Phi(\mathbf{x}_i\boldsymbol{\beta})$ ,
- Logit  $F(\mathbf{x}_i) = \Lambda(\mathbf{x}_i\boldsymbol{\beta})$ .

gdzie  $\Phi(\cdot)$  jest dystrybuantą rozkładu standardowego Normalnego  $N(0, 1)$ ,  
a  $\Lambda(\cdot)$  jest dystrybuantą rozkładu logistycznego postaci:

$$\frac{e^{x_i\beta}}{1 + e^{x_i\beta}}$$

Po oszacowaniu modeli ogólnych, dokonany został wybór między logitem i probitem na podstawie wartości kryteriów informacyjnych Akaike i Bayesa-Schwarza (LPM jest odrzucany ze względu na możliwość otrzymania dopasowanych prawdopodobieństw nienależących do przedziału  $[0; 1]$ ). Następnie na wybranym modelu zastosowana została procedura od ogółu do szczegółu przy wykorzystaniu testu ilorazu wiarygodności. Ostatnim krokiem było sprawdzenie dopasowania modelu za pomocą testu typu związku (*linktest*) będącego uogólnieniem testu RESET Ramsey'a, co skutkowało koniecznością zmiany formy funkcyjnej modelu. Model ogólny posłużył do uzyskania tabeli efektów cząstkowych (bowiem lepszy okazał się model probitowy) oraz przetestowania hipotez, w tym celu dodana została do niego zmienna Altmana. Następnie zaprezentowane zostały krzywe ROC wraz z punktacją AUC-score dla obu finalnych modeli oraz wybrany został optymalny punkt odcięcia klasyfikacji bankrutów i nie bankrutów ( $p^*$ ), tak aby uzyskać model odpowiadający oczekiwaniom stawianym w głównej hipotezie badawczej.

I hipoteza poboczna została przetestowana za pomocą formuły zawartej we wstępie, tj. zbadaniu efektu cząstkowego dla średnich charakterystyk dla zmiennej Rozmiar.

II hipoteza badawcza zweryfikowana została przy zastosowaniu zaprezentowanej poniżej procedury opartej na symulacji bootstrapowej (Bi-Huei Tsai (2013)) [10]:

1. odrębne losowanie ze zwracaniem bankrutów i nie bankrutów w proporcjach z oryginalnej próby,
2. oszacowanie modelu z i bez zmiennej Altmana (probit 13 i probit 15),
3. arbitralny wybór progu odcięcia i na jego podstawie dobór  $\theta_1$  i  $\theta_0$  dla których jest on optymalny,
4. obliczenie  $p(\hat{y}_i < p^* | y_i = 1)$  oraz  $p(\hat{y}_i > p^* | y_i = 0)$  dla wylosowanego zbioru danych, wyestymowanych modeli i progów odcięcia,
5. obliczenie i przeprowadzenie jednostronnego testu na równość średnich  $EMC$  dla każdego progu odcięcia, przy hipotezie alternatywnej  $EMC_{standard} > EMC_{Altman}$ , procedura została powtórzona 1000 razy dla każdego progu odcięcia z przedziału 0.01 - 0.5.

$EMC$  obliczone zostało przy założeniu, że punkt odcięcia wybierany jest w oparciu o ekonomiczne koszty błędu klasyfikacji i jest optymalny. Clive, Granger i Pesaran (2000) za optymalny punkt

odcięcia ( $p^*$ ) wskazują

$$p^* = \frac{1}{1 + \frac{\theta_1}{\theta_0}}$$

znając zatem  $p^*$  można obliczyć stosunek  $\theta_1$  i  $\theta_0$ , tym samym umożliwiając obliczenie *EMC* [9].

Dla wszystkich testów przyjęto próg istotności  $\alpha = 0.05$

## 5 Wyniki

Tabela 4 przedstawia zbiorcze wyniki estymacji modeli ogólnych (LPM, Probit ogólny, Logit ogólny) oraz finalnych (Probit 13 - wynik procedury od ogółu do szczegółu bez zmiennej Altmana, Probit 15 - Probit 13 uwzględniający zmienną Altmana). Czytelnik może zastanawiać się nad oszacowaniem parametru dla kwadratu zmiennej Kapitał obrotowy/Aktywa, bowiem interakcji tej nie było w spisie zmiennych poddawanych analizie prezentowanych w Tabeli 1 oraz 2. Pojawienie się wspomnianej potęgi było wynikiem przeprowadzenia testu typu związku. Surowy model wynikowy z procedury od ogółu do szczegółu (Probit 12) nie odrzucał hipotezy zerowej o nieistotności interakcji zmiennych egzogenicznych, odnaleziono zatem brakujący element i po powtórным przeprowadzeniu testu, hipoteza zerowa została odrzucona, co oznaczało uzyskanie poprawnej formy funkcyjnej modelu. Ponadto test ilorazu wiarygodności wskazał na odrzucenie hipotezy zerowej o nieistotności kwadratu zmiennej Kapitał obrotowy/Aktywa ( $p\_value = 0.0004$ , Probit 12 to model bazowy), na podstawie czego można wnioskować o przydatności zmiennej w modelu. Poprawną formę funkcyjną stwierdzono również dla modelu Probit 15, zmienna Altmana nie okazała się jednak istotna statystycznie. W Tabeli 3 przedstawiono wyniki testów typu związku. Warto odnotować, że w przypadku estymacji LPM program zwrócił informację o możliwym wystąpieniu zjawiska niedokładnej współliniowości, wskaźniki inflacji wariacji nie wskazały jednak winnego.

Tablica 3: wyniki testów typu związku

	<b>Probit 12</b>	<b>Probit 13</b>	<b>Probit 15</b>
<b>Stała</b>	-0.1602** (0.0715)	-0.0653 (0.0823)	-0.0802 (0.0803)
<b>yhat</b>	0.4347*** (0.0410)	0.5773*** (0.0644)	0.5553*** (0.0621)
<b>yhat2</b>	-0.0171** (0.0086)	0.0164 (0.0130)	0.0114 (0.0128)

Źródło: Opracowanie własne

Tablica 4: Wyniki estymacji

	LPM	Probit ogólny	Logit ogólny	Probit 13	Probit 15
Stała	0.0817 (0.0792)	-2.2617*** (0.5473)	-5.2161*** (1.1555)	-1.3049*** (0.1094)	-1.4493*** (0.1426)
Current ratio	0.0004 (0.0003)	0.0032 (0.0021)	0.0058 (0.0036)		
Dźwignia finansowa	-0.0000 (0.0000)	-0.0002 (0.0009)	-0.0005 (0.0019)		
Kapitał obrotowy/Aktywa	-0.0640*** (0.0213)	-0.4542*** (0.1405)	-1.0104*** (0.2831)	-0.7652*** (0.1093)	-0.7384*** (0.1106)
I(Kapitał obrotowy/Aktywa ** 2)				-0.1470*** (0.0355)	-0.1540*** (0.0354)
Rozmiar	-0.0013 (0.0316)	0.4060 (0.2644)	1.2552** (0.5595)		
I(Rozmiar ** 2)	-0.0023 (0.0034)	-0.0761** (0.0338)	-0.2048*** (0.0720)	-0.0308*** (0.0054)	-0.0275*** (0.0058)
Marża zysku brutto	-0.0541** (0.0256)	-0.3611*** (0.1339)	-0.6383*** (0.2477)	-0.3078** (0.1313)	-0.3213** (0.1315)
Przyrost sprzedaży - spadek/brak danych	0.0908*** (0.0099)	0.7873*** (0.0729)	1.6757*** (0.1557)	0.7991*** (0.0695)	0.8046*** (0.0696)
ROA	-0.1113*** (0.0313)	-0.5996*** (0.1812)	-1.2652*** (0.4059)	-0.6446*** (0.1753)	-0.6327*** (0.1752)
Rotacja należności	-0.0002* (0.0001)	-0.0012* (0.0006)	-0.0017 (0.0013)		
Rotacja zapasów	-0.0001 (0.0001)	-0.0007 (0.0007)	-0.0010 (0.0013)		
Rotacja zobowiązań	0.0003*** (0.0001)	0.0012*** (0.0004)	0.0018*** (0.0007)	0.0006** (0.0003)	0.0007** (0.0003)
Sprzedaż/Aktywa	0.0087 (0.0054)	0.0292 (0.0285)	0.0920 (0.0576)		
Wrt. rynkowa/Wrt. księgowa długu	-0.0000 (0.0000)	-0.0019 (0.0024)	-0.0028 (0.0042)		
Zobowiązania/Aktywa	-0.0052 (0.0156)	0.0468 (0.1156)	0.0781 (0.2088)		
Zyski zatrzymane/Aktywa	-0.0008 (0.0007)	-0.0111 (0.0204)	-0.0194 (0.0406)		
wypłacalność	-0.0411 (0.0291)	-0.1391 (0.2039)	-0.2392 (0.4293)		
wypłacalność:Przyrost sprzedaży - spadek/brak danych	0.1998*** (0.0569)	0.1652 (0.2254)	0.1157 (0.4426)		
Zscore					0.0408 (0.0257)
$PseudoR^2(R^2 dla LPM)$	0.162	0.233	0.231	0.234	0.235
$R^2_{Count}$	-	0.9384	0.938	0.938	0.9373
$R^2_{McKelvey\&Zavoina}$	-	0.027	0.0203	0.0203	0.0101
AIC	-744.6922	1727.69	1733.01	1706.12	1705.69
BIC	-628.5925	1843.79	1849.11	1757.72	1763.75

Oszacowania parametrów dla modelu finalnego - Probit 13 zgodne są z intuicją i literaturą, wyjątkiem jest tutaj zmienna Rotacja zobowiązań, bowiem w artykule Apostolosa i in. (2019) przyjęła znak -" [3]. Rotacja zobowiązań ma jednak dwuliniowy charakter, z jednej strony może świadczyć o problemach z płynnością, z strony może być oznaką mocnej pozycji i dogodnych umów handlowych, pozwalających na regulację zobowiązań w dłuższych interwałach. Sprzeczny zarówno z intuicją, jak i literaturą (Altman (1968), Bi-Huei Tsai (2013)) okazuje się znak oszacowania parametru Zscore - model probit 15 nie służy jednak do wnioskowania statystycznego [2] [10].

Zaprezentowane w Tabeli 4 współczynniki Pseudo  $R^2$  zostały policzone dla standardowego progu odcięcia - 0.5. Można z nich odczytać, że ogólny model probitowy charakteryzował się największą skutecznością - 93.84% obserwacji zostało poprawnie zaklasyfikowanych, modele ogólny logitowy i probit 13 zanotowały taką samą skuteczność (93.8%), najgorzej natomiast wypadł model probit 15 (93.73%). Dopasowany  $R^2$  liczebnościowy mówi z kolei o procencie poprawnych klasyfikacji wynikających ze zmienności zmiennych egzogenicznych zawartych w modelach, pod tym względem również najlepszy okazał się ogólny model probitowy (2.27%), modele ogólny logitowy i probit 13 ponownie uzyskały identyczny wynik (2.03%), najgorzej wypadł model probit 15 (1.01%).  $R^2$  McKelveya i Zavoina szereguje modele w nieco innej kolejności i mówi o procencie wyjaśnialności zmienności zmiennej ukrytej za pomocą modelu, gdyby ta była obserwowalna, najlepszy okazał się ogólny logit (33.03%), potem ogólny probit (28.47%), probit 13 (27.7%) i probit 15 (27.57%). Pod nazwą Pseudo  $R^2$  kryje się  $R^2$  McFaddena i posiada on interpretacji.

W Tabeli 5 przedstawione zostały efekty cząstkowe dla średnich charakterystyk z modelu Probit 13 (wynik procedury od ogółu do szczegółu z korektą formy funkcyjnej, nieuwzględniający zmiennej Altmana czyli finalny model do wnioskowania statystycznego (zmienna Altmana była nieistotna)).

Tablica 5: Efekty cząstkowe Probitu 13

<b>Zmienna zależna:</b>	bankrut					
<b>Metoda:</b>	charakterystyki średnie					
	dy/dx	std err	z	P>  z	[0.025	0.975]
<b>I(Kapitał obrotowy/Aktywa ** 2)</b>	-0.0111	0.003	-4.213	0.000	-0.016	-0.006
<b>I(Rozmiar ** 2)</b>	-0.0023	0.000	-5.704	0.000	-0.003	-0.002
<b>Kapitał obrotowy/Aktywa</b>	-0.0577	0.008	-6.907	0.000	-0.074	-0.041
<b>Marża zysku brutto</b>	-0.0232	0.010	-2.330	0.020	-0.043	-0.004
<b>Przyrost sprzedaży - spadek/brak danych</b>	0.0858	0.009	9.047	0.000	0.067	0.104
<b>ROA</b>	-0.0486	0.013	-3.607	0.000	-0.075	-0.022
<b>Rotacja zobowiązań</b>	4.319e-05	2.11e-05	2.050	0.040	1.9e-06	8.45e-05

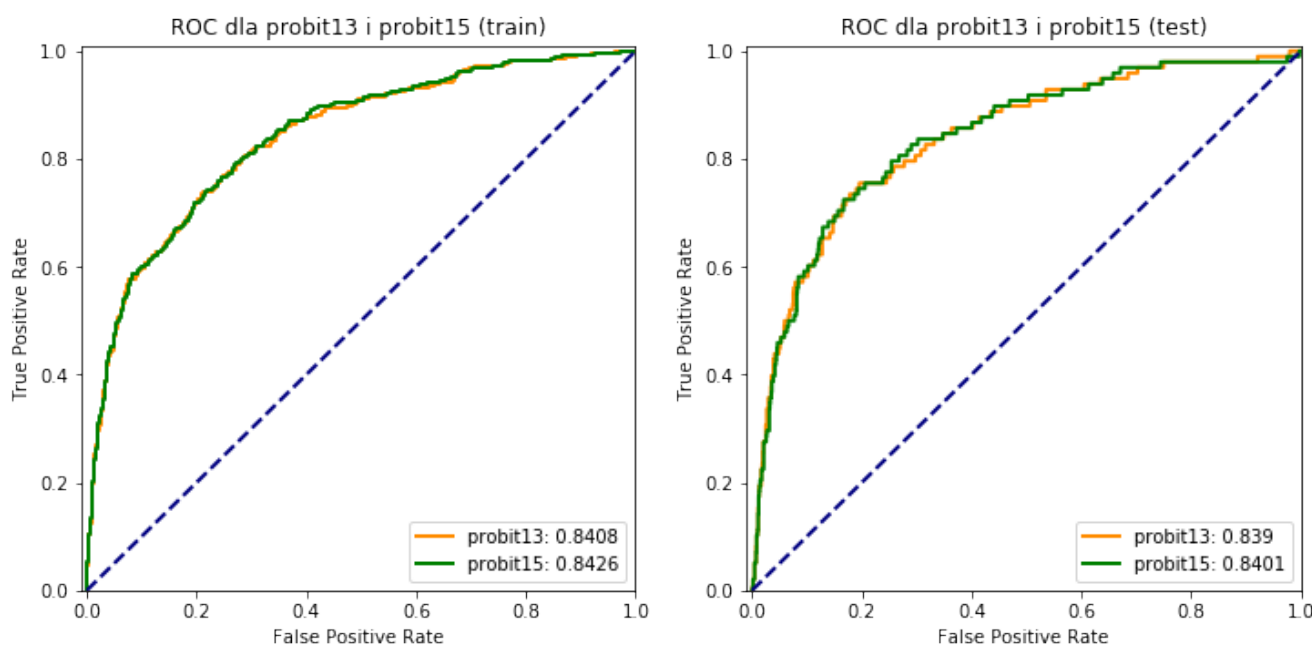
Źródło: Opracowanie własne

Z uwagi na fakt nieinterpretowalności wielkości parametrów modelu, a jedynie ich znaków, przygotowano powyższą tabelę efektów cząstkowych, można się z niej dowiedzieć, że dla średnich charakterystyk, zmiennymi wpływającymi na zwiększenie prawdopodobieństwa bankructwa jest Spadek sprzedaży w stosunku do roku ubiegłego lub brak danych, występowanie tej zmiennej wiąże się ze wzrostem p. klasyfikacji jako bankruta o 8.58 p. procentowych, dużo słabiej oddziałuje Rotacja zobowiązań - jeżeli wydłuży się o jeden dzień, prawdopodobieństwo bankructwa wzrośnie o 0.0004 p. procentowego. Wzrost pozostałych wskaźników obniżają prawdopodobieństwo bankructwa. Najsilniej wpływa stosunek Kapitału obrotowego do Aktywów, bowiem wzrostowi wskaźnika o jednostkę towarzyszy spadek zmiennej zależnej o 5.31 p. procentowych, w przypadku ROA jest to 4.86 p. procentowych, dużo słabiej oddziałuje Marża zysku brutto (2.32 p. procentowych) i Rozmiar (0.96 p.). Określenie ilorazu szans możliwe jest tylko dla modelu logitowego, co jest jego dużą zaletą, bowiem można wtedy interpretować model w odniesieniu do wszystkich obserwacji, a nie tylko wybranych charakterystyk (w niniejszej pracy - średnich, jednak podobne wyniki uzyskano dla uśrednienia wszystkich efektów cząstkowych).

Zwróciwszy uwagę na efekt cząstkowy Rozmiaru oraz znak przy oszacowaniu modeli probitowych można stwierdzić, że nie ma podstaw do odrzucenia I pobocznej hipotezy badawczej.

Poniżej za pomocą krzywych ROC przedstawiono jakość modeli. Dla przypomnienia, najlepszy model Charalambakisa i Garretta (2019) punktacja ROC-AUC wyniosła 0.8191, a dla modelu Christopoulou i in. (2019) 0.7377.

Rysunek 5: Krzywe ROC dla probit 13 i probit 15



Źródło: Opracowanie własne.

Na wykresach mamy okazję zaobserwować bardzo dobrą generalizację modeli probit 13 i 15 (z symboliczną przewagą tego drugiego), bowiem wyniki na zbiorach testowych i treningowych różnią się maksymalnie o około 0.025, przy ogólnej punktacji w granicach 0.839 - 0.8426, co w porównaniu z estymatorem losowym, którego wynik z definicji wynosi 0.5 (przerywana linia 45 stopni na wykresach) świadczy o tym, że modele mają wysoki potencjał predykcyjny.

Za optymalny punkt odcięcia, tj. najbardziej odpowiadający oczekiwaniom (jak najwyższa wrażliwość przy specyficzności co najmniej 0.95) uznano punkt  $p^* = 0.19$ . Wybrany on został na podstawie analizy wyników na zbiorze treningowym dla punktów odcięcia z przedziału 0.01 - 0.5. Poniżej zaprezentowano tablice klasyfikacji i predykcji dla standardowego (Tabela 6) i optymalnego (Tabela 7) punktu odcięcia.

Tablica 6: Tablice klasyfikacyjne i predykcyjne probit 13 i 15 ( $p^* = 0.5$ )

Probit 13, cutoff 0.5 (train)				Probit 13, cutoff 0.5 (test)			
		Obserwowane				Obserwowane	
		0	1			0	1
Przewidywane	0	4350	261	Przewidywane	0	1066	82
	1	29	35		1	13	16
Probit 15, cutoff 0.5 (train)				Probit 15, cutoff 0.5 (test)			
		Obserwowane				Obserwowane	
		0	1			0	1
Przewidywane	0	4348	262	Przewidywane	0	1066	82
	1	31	34		1	13	16

Źródło: Opracowanie własne.

Tablica 7: Tablice klasyfikacyjne i predykcyjne probit 13 i 15 ( $p^* = 0.19$ )

Probit 13, cutoff 0.19 (train)				Probit 13, cutoff 0.19 (test)			
		Obserwowane				Obserwowane	
Przewidywane		0	1	Przewidywane		0	1
	0	4161	161		0	1024	53
	1	218	135		1	55	45
Probit 15, cutoff 0.19 (train)				Probit 15, cutoff 0.19 (test)			
		Obserwowane				Obserwowane	
Przewidywane		0	1	Przewidywane		0	1
	0	4170	162		0	1028	53
	1	209	134		1	51	45

Źródło: Opracowanie własne.

Z tak przygotowanych tablic można łatwo obliczyć specyficzność  $[P(\hat{p} < p^* | y_i = 0)]$  i wrażliwość  $[P(\hat{p} > p^* | y_i = 1)]$  modeli na każdej z prób. Dla standardowego progu odcięcia 0.5 model probit 13 zanotował specyficzność 0.9934 (próba treningowa) i 0.988 (próba testowa) oraz wrażliwość 0.1182 (próba treningowa) i 0.1633 (próba testowa). Model probit 15 uzyskał specyficzność 0.9929 (próba treningowa) i 0.988 (próba testowa) oraz wrażliwość 0.1149 (próba treningowa) i 0.1633 (próba testowa).

Dla optymalnego progu odcięcia probit 13 to specyficzność 0.9502 (próba treningowa) i 0.949 (próba testowa) oraz wrażliwość 0.4561 (próba treningowa) i 0.4592 (próba testowa). Model probit 15 z optymalnym progiem odcięcia charakteryzuje się specyficznością 0.9523 (próba treningowa) i 0.9527 (próba testowa) oraz wrażliwością 0.4527 (próba treningowa) i 0.4592 (próba testowa). Skuteczność modelu probit 13 to 91.9% na próbie treningowej i 91% na próbie testowej, a modelu probit 15 to 92% na próbie treningowej i 91.16% na próbie testowej.

W tym miejscu można obronić główną hipotezę badawczą - za pomocą jedynie sprawozdania finansowego jesteśmy w stanie z wysoką skutecznością odróżnić bankrutów od nie bankrutów (w horyzoncie rocznym). Hipoteza w swoim pierwotnym brzmieniu zakładała stworzenie systemu wczesnej informacji, co oznacza, że model przewidując nie bankruta daje przynajmniej 95% pewności (wybór na podstawie standardowego progu istotności przyjętego w statystyce), przy wykryciu przynajmniej 1/3 liczby bankrutów. Po doborze optymalnego poziomu odcięcia (0.19) udało się uzyskać model spełniający założenia o specyficzności jednocześnie wykrywając znacznie większą liczbę bankrutów (45 - 46%), będąc jednocześnie nieprzetrenowanym (*not overfitted*) - wyniki stabilne niezależnie od próby (treningowa, testowa). Dokładne wyniki bootstrapowej symulacji



mającej za zadanie zweryfikowanie wpływu dodania zmiennej Altmana zaprezentowano w Tabeli 8.

Tablica 8: Wyniki bootstrapu

cutoff	U	wartość krytyczna	p_value	cutoff	U	wartość krytyczna	p_value
0.01	-0.1303	1.6449	0.4482	0.26	1.7424	1.6449	0.0407
0.02	3.8996	1.6449	0.0	0.27	1.4229	1.6449	0.0774
0.03	6.112	1.6449	0.0	0.28	1.467	1.6449	0.0712
0.04	2.4671	1.6449	0.0068	0.29	1.9201	1.6449	0.0274
0.05	2.5763	1.6449	0.005	0.3	1.8985	1.6449	0.0288
0.06	1.2859	1.6449	0.0992	0.31	2.1166	1.6449	0.0171
0.07	0.3911	1.6449	0.3479	0.32	1.4535	1.6449	0.073
0.08	-0.1501	1.6449	0.4404	0.33	1.1468	1.6449	0.1257
0.09	-0.597	1.6449	0.2752	0.34	0.575	1.6449	0.2826
0.1	0.4698	1.6449	0.3192	0.35	0.369	1.6449	0.3561
0.11	0.6641	1.6449	0.2533	0.36	0.257	1.6449	0.3986
0.12	0.4529	1.6449	0.3253	0.37	-0.1895	1.6449	0.4248
0.13	2.0259	1.6449	0.0214	0.38	0.0088	1.6449	0.4965
0.14	3.629	1.6449	0.0001	0.39	0.0398	1.6449	0.4841
0.15	2.1689	1.6449	0.015	0.4	-0.2804	1.6449	0.3896
0.16	1.0468	1.6449	0.1476	0.41	-0.0335	1.6449	0.4866
0.17	-0.0196	1.6449	0.4922	0.42	0.0788	1.6449	0.4686
0.18	0.4998	1.6449	0.3086	0.43	-0.519	1.6449	0.3019
0.19	1.3758	1.6449	0.0844	0.44	-0.3427	1.6449	0.3659
0.2	1.5945	1.6449	0.0554	0.45	-0.5609	1.6449	0.2874
0.21	0.9782	1.6449	0.164	0.46	-0.6426	1.6449	0.2602
0.22	0.9225	1.6449	0.1781	0.47	-0.8392	1.6449	0.2007
0.23	0.6066	1.6449	0.2721	0.48	-0.9368	1.6449	0.1744
0.24	0.4947	1.6449	0.3104	0.49	-1.0984	1.6449	0.136
0.25	1.3794	1.6449	0.0839	0.5	-1.5959	1.6449	0.0553

Źródło: Opracowanie własne.

Hipoteza zerowa przeprowadzonego w Tabeli 8 testu mówi o równości średnich, zatem II poboczna hipoteza badawcza zostaje obroniona w przypadku odrzucenia hipotezy zerowej testu statystycznego (hipoteza alternatywna testu to  $\overline{EMC}_{probit13} > \overline{EMC}_{probit15}$ ). Opisany scenariusz miał miejsce w przypadkach progów 0.02, 0.03, 0.4, 0.5, 0.13, 0.14, 0.15, 0.26, 0.29, 0.3 oraz 0.31, co daje w sumie 11 na 50 progów odcięcia, w których  $EMC$  okazał się niższy na poziomie istotności 5% dla modelu zawierającego zmienną Altmana od modelu bez tej zmiennej. Z jednej

strony proporcja 11/50 wskazuje, że w większości przypadków zmienna Altmana jest bezużyteczna w wyestymowanym modelu, z drugiej strony dla konkretnych przypadków  $\theta_0$  i  $\theta_1$  może okazać się dobrym rozwiązaniem. Zatem to czy II poboczna hipoteza badawcza zostanie przyjęta, tak jak wiele rzeczy w ekonomii, zależy od sytuacji.

## 6 Zakończenie

Badanie bankructwa przedsiębiorstw nie należy do spraw łatwych, po pierwsze ze względu na niejednolite przepisy sprawozdawcze, po drugie ze względu na dopuszczalną "plastyczność" dokumentów finansowych, po trzecie na możliwe definiowanie bankructwa w różny sposób (momenty w czasie życia przedsiębiorstwa) i po czwarte ze względu na silne niebilansowanie populacji (w próbie wykorzystanej w niniejszym badaniu proporcja to około 6/94). Ostatnie wynika z rzadkości obserwowania zdarzenia, co z punktu widzenia gospodarki jest dobrą oznaką.

W niniejszej pracy udało się stworzyć model o możliwościach przekraczających stawiane przed nim oczekiwania, nie jest to jednak najlepszy model tego typu, bowiem w literaturze można znaleźć przykłady znacznie większej wykrywalności "1", przy zachowaniu zbliżonej wykrywalności "0" oraz skuteczności, jak chociażby modele Ohlsona (1980) czy Yazdanfar (2011) [8] [12]. Publikowane były również modele z gorszymi (Ni (2013)) i podobnymi (Waqas i Md-Rus (2018)) wynikami [6] [5]. Ogólne wyniki są zatem więcej niż zadowalające. Ponadto okazało się, że po odpowiednim przygotowaniu bazy można z powodzeniem wykorzystywać informację o braku danych (podejście zaproponowane przez Zmijewskiego (1984)) [13].

Zweryfikowano również slogan "too big to fail", a wyniki nie pozwoliły na uznanie jego nieprawdziwości. W badanym przypadku nie udało się jednoznacznie odpowiedzieć na pytanie zadane przez Bi-Huei Tsai (2013) o poprawie modelu po dodaniu do niego punktacji Z-score, odpowiedź okazała się zależna od ekonomicznych kosztów błędu klasyfikacji zdefiniowanych przez badacza [10].

W związku z rozwojem technologii bazujących na Sztucznej Inteligencji i Uczeniu Maszynowym, dalsze prace nad tematem będą dotyczyć próby skonstuowania skuteczniejszych modeli przy wykorzystaniu tych metod.

## Literatura

- [1] Ustawa z dnia 28 lutego 2003 prawo upadłościowe. Art. 10, Art 11 ust. 1, Art 21 ust. 1.
- [2] Edward I. Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4):589–609, 1968. Impact factor: 6.2.

- [3] Petros Kalantonis Theodora Koukkou Apostolos G. Christopoulos, Ioannis G. Dokas. Investigation of financial distress with a dynamic logit based on the linkage between liquidity and profitability status of listed firms. *Journal of the Operational Research Society*, 70(10):1817–1829, 2019. Impact factor: 1.077.
- [4] Ian Garrett Evangelos C. Charalambakis. On corporate financial distress prediction: What can we learn from private firms in a developing economy? evidence from greece. *Review of Quantitative Finance and Accounting*, 52:467–491, 2019. Impact factor: 1.66.
- [5] Rohani Md-Rus Hamid Waqas. Predicting financial distress: Applicability of o-score and logit model for pakistani firms. *Business and Economic Horizons*, 14(2):389–401, 2018. Impact factor: 0.12.
- [6] Xiaoyan Cheng Guan Gong Jinlan Ni, Wikil Kwak. The determinants of bankruptcy for chinese firms. *Review of Pacific Basin Financial Markets and Policies*, 17(2), 2014. Impact factor: 0.24.
- [7] Jakub M. Tomczak Maciej Zieba, Sebastian K. Tomczak. Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Systems With Applications*, 58(1):93–101, 2016. Impact factor: 4.292.
- [8] James A. Ohlson. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18(1):101–131, 1980. Impact factor: 4.891.
- [9] Clive W. J. Granger M. Hashem Pesaran. Economic and statistical measures of forecast accuracy. *Journal of Forecasting*, 19(7):537–560, 2000. Impact factor: 0.934.
- [10] Bi-Huei Tsai. An early warning system of financial distress using multinomial logit models and a bootstrapping approach. *Emerging Markets Finance Trade*, 49(2):43–69, 2013. Impact factor: 0.60.
- [11] Jacek Welc. Ebitda vs. cash flows in bankruptcy prediction on the polish capital market. *European Financial and Accounting Journal*, 12(2):91–103, 2017. Impact factor: 0.898 (2015).
- [12] Darush Yazdanfar. Predicting bankruptcy among smes: evidence from swedish firm-level data. *International Journal of Entrepreneurship and Small Business*, 14(4):551–565, 2011. Impact factor: 0.42.
- [13] Mark E. Zmijewski. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22:59–82, 1984. Impact factor: 4.891.