

Potencjał sztucznej inteligencji w prognozowaniu upadłości spółek – analiza porównawcza dyskryminacyjnych modeli predykcji bankructwa polskich przedsiębiorstw

Debora Kotagomba

Uniwersytet Ekonomiczny we Wrocławiu

e-mail: 191895@student.ue.wroc.pl

ORCID: [0009-0002-3629-4525](https://orcid.org/0009-0002-3629-4525)

Kacper Kobylański

Uniwersytet Ekonomiczny we Wrocławiu

e-mail: 191868@student.ue.wroc.pl

ORCID: [0009-0007-3807-0927](https://orcid.org/0009-0007-3807-0927)

Jakub Petyniak

Uniwersytet Ekonomiczny we Wrocławiu

e-mail: 191871@student.ue.wroc.pl

ORCID: [0009-0005-0131-2289](https://orcid.org/0009-0005-0131-2289)

© 2025 Debora Kotagomba, Kacper Kobylański, Jakub Petyniak

Praca opublikowana na licencji Creative Commons Uznanie autorstwa-Na tych samych warunkach 4.0 Międzynarodowe (CC BY-SA 4.0). Skrócona treść licencji na <https://creativecommons.org/licenses/by-sa/4.0/deed.pl>

Cytuj jako: Kotagomba, D., Kobylański, K. i Petyniak, J. (2025). Potencjał sztucznej inteligencji w prognozowaniu upadłości spółek – analiza porównawcza dyskryminacyjnych modeli predykcji bankructwa polskich przedsiębiorstw. W: M. Gasz, A. Koza, A. Politał (red.), *Ekonomia i finanse* (s. 44-58). Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu.

Streszczenie: Przewidywanie ryzyka upadłości przedsiębiorstw to kluczowe wyzwanie dla menedżerów i instytucji finansowych. Dokładna ocena zagrożenia umożliwia wdrożenie odpowiednich działań zapobiegawczych. Istnieje wiele modeli predykcji upadłości opartych na analizie finansowej, dyskryminacyjnej czy metodach statystycznych. W tym opracowaniu postanowiono porównać klasyczne, polskie modele dyskryminacyjne (m.in. Gajdki i Stosa) z nowoczesnymi rozwiązaniami zaproponowanymi

przez ChatGPT – potocznie nazywany „sztuczną inteligencją” – model LLM (*large language model*), pod względem skuteczności ich predykcji. Celem niniejszego rozdziału jest sprawdzenie w praktyce owych narzędzi, a także zbadanie ich potencjału oraz określenie, czy mogą znaleźć zastosowanie w świecie zarządzania finansami. Analiza wykazała, że systemy wczesnego ostrzegania wykreowane przez SI mogą być skuteczniejsze od niektórych tradycyjnych modeli i wykazują pewien potencjał w prognozowaniu ryzyka bankructwa. Badania autorów niniejszego rozdziału zostały przeprowadzone na wybranych upadłych spółkach oraz opierały się na metodzie analizy porównawczej, analizie danych i sprawozdań finansowych, krytycznej analizie literatury, ocenie efektywności poszczególnych modeli predykcji upadłości oraz wykorzystaniu modeli językowych.

Słowa kluczowe: modele predykcji upadłości, ChatGPT, upadłość przedsiębiorstw, analiza dyskryminacyjna

1. Geneza i istota dyskryminacyjnych modeli predykcji upadłości

Analiza dyskryminacyjna zaczęła być stosowana w prognozowaniu upadłości firm w latach 60. XX wieku. Kluczowe znaczenie miały badania Altmana z 1968 roku, w których opracował model klasyfikacyjny oparty na pięciu wskaźnikach finansowych. Jego prace stały się fundamentem dla dalszego rozwoju tej metodologii i zyskały szerokie uznanie. Kolejne prace były kontynuowane przez licznych autorów, którzy opracowywali modele klasyfikacyjne dostosowane do specyfiki gospodarek różnych krajów, a także wykorzystywali coraz bardziej zaawansowane metody wielowymiarowej analizy danych.

Altman, tworząc tzw. Z-score, po raz pierwszy połączył dane finansowe z metodą analizy dyskryminacyjnej, co stanowiło przełom w podejściu do przewidywania upadłości. Od tego czasu analiza dyskryminacyjna była szeroko stosowana w różnych krajach i branżach w celu opracowania efektywnych narzędzi diagnostycznych, pozwalających na ocenę kondycji finansowej firm i prognozowanie ich przyszłych losów (Kisielińska i Waszkowski, 2010).

Systemy wczesnego ostrzegania przed upadłością to stosunkowo nowa i niedoceniana dziedzina. W Polsce rośnie jej znaczenie, jednak zarządy firm zwykle wdrażają nowoczesne metody dopiero po utracie płynności finansowej, a zatrudnienie analityka finansowego często świadczy o kryzysie firmy. Mimo tego opinie ekspertów są często ignorowane, a same systemy wczesnego ostrzegania lekceważone. Wdrożenie takich systemów wymaga, mimo pozorów prostoty, dużego doświadczenia i intuicji, ponieważ zastosowanie niewłaściwego modelu może prowadzić do wyników całkowicie odbiegających od rzeczywistości (Kisielińska i Waszkowski, 2010).

Modele predykcji upadłości oceniają kondycję finansową firm, identyfikując sygnały pogarszającej się płynności i ryzyka bankructwa. Stanowią one wyłącznie narzędzie do wstępnej analizy, bowiem wymagają monitorowania i nie oferują rozwiązań naprawczych. Często utożsamia się je z systemami prognozującymi upadłość, ponieważ ich głównym celem jest wykrywanie firm zagrożonych bankructwem. W literaturze znaleźć można wiele takich modeli przedsiębiorstw dostosowanych do realiów polskiego rynku. Znakomitą ich większość stanowią modele oparte na

liniowej analizie dyskryminacyjnej. Oznacza to w dużym uproszczeniu, że model taki jest sumą współczynników funkcji dyskryminacyjnej pomnożonych przez wektor danej cechy. Na podstawie wartości wyniku równania określana jest klasa przynależności danego przedsiębiorstwa, tzn. jeśli liczba ta jest powyżej określonej wartości, to przedsiębiorstwo należy do grupy A, jeśli poniżej, to należy do grupy B (Tłuczak, 2013).

W przypadku modeli predykcji upadłości wyróżnia się następujące klasy: przedsiębiorstwa w dobrej kondycji oraz takie, które zagrożone są bankructwem. Często także występuje tzw. szara strefa wyników, w której skuteczność predykcji jest znacząco niższa niż w pozostałych przypadkach. Istnieją również modele oparte na funkcji logitowej (nazywanej też logistyczną), jednak nie zostaną one poruszone przez autorów niniejszej pracy mimo banalnej interpretacji otrzymywanych wyników i dużej poprawności klasyfikacji (Piszczek, 2009).

Należy również wspomnieć, że obecnie używane modele nie są wolne od różnych wad, z których najczęściej wymieniane to:

- ograniczona skuteczność z powodu wyselekcjonowania danych,
- brak uwzględnienia pozafinansowych czynników, takich jak [m.in.](#) sytuacja makroekonomiczna,
- konieczność stałej aktualizacji wskaźników (Banach-Kobyra, 2017).

2. Wstęp do badań

Modele predykcji upadłości przedsiębiorstw są ważnym narzędziem zarówno dla menedżerów czy instytucji finansowych, jak i dla inwestorów. Precyzyjna ocena ryzyka upadłości umożliwia podjęcie odpowiednich działań – wdrożenie działań zapobiegawczych lub identyfikację przyczyn problemu.

W ocenie kondycji finansowej jednym z pomocnych narzędzi jest analiza wskaźnikowa, badająca relacje rachunkowo-finansowe w czasie. Są bardziej skomplikowane przypadki, takie jak przedsiębiorstwa wysoko rentowne, ale bardzo zadłużone lub z niską płynnością (Wojnar, 2014, s. 220). W zarządzaniu przedsiębiorstwem niezbędne staje się więc stosowanie szerokiego wachlarza starannie dopasowanych narzędzi umożliwiających wczesne wykrycie ryzyka upadłości.

Stworzenie skutecznego modelu predykcji wymaga eksperckiej wiedzy z zakresu finansów czy mikroekonomii. Obecnie jednak narzędzia analizujące źródła ułatwiają dostęp do kluczowych informacji, częściowo zastępując wieloletnie zdobywanie wiedzy. Jednym z takich narzędzi jest ChatGPT – potocznie nazywany Sztuczną Inteligencją – LLM (*large language model*) stworzony przez OpenAI. Potrafi on przeszukiwać pliki lub internet oraz generować logicznie spójne odpowiedzi na pytania użytkowników. Technologia ta opiera się na modelu typu transformer (Kocoń i in., 2023) – tekst wprowadzony przez użytkownika jest zamieniany na kilkunastokrotne tokeny, analizowany względem znanych programowi wzorców, a następnie określana

jest „ważność” poszczególnych słów, co pozwala modelowi „zrozumieć” sens wypowiedzi i wygenerować odpowiedź. ChatGPT nie myśli samodzielnie, a jego działanie przypomina wyszukiwarkę analizującą dane i identyfikującą schematy (Vaswani i in., 2017). Może być przydatnym narzędziem do analizy danych, jednak wymaga nauki w bardziej niszowych obszarach. Najlepszym tego przykładem jest seria filmów autorstwa Levy’ego Rozmana, znanego szerzej jako GothamChess, który na przełomie kilku miesięcy podejmował próby rozgrywania z programem partii szachowych. Początkowe rozgrywki były abstrakcyjne, jednak po kilku miesiącach trenowania modelu był on w stanie rozegrać znacznie bardziej zgodną z zasadami partię, ponieważ zaczął coraz mocniej je rozumieć.

Autorzy monografii analizują, czy ChatGPT potrafi tworzyć skuteczne modele przewidujące upadłość przedsiębiorstw oraz oceniające ich efektywność. Modele te porównano z powszechnie stosowanymi systemami prognozowania upadłości dostosowanymi do polskiego rynku. Za materiał empiryczny posłużyły dane firm, wobec których ogłoszono upadłość, uznając za trafne prognozy, które przewidywały bankructwo. Szczegółowe kryteria oceny modeli zostaną omówione w dalszej części rozdziału.

Badania wykonywane były na danych znanych, upadłych spółek, zebranych przy użyciu oprogramowania EMIS Next oraz w środowisku programu Excel.

3. Wstępna analiza modeli i ich selekcja

Pierwszym etapem pracy było wybranie odpowiednich przedsiębiorstw, na danych których opierać się miało badanie skuteczności predykcji wybranych do analizy systemów wczesnego ostrzegania. Spółkami, które zostały uznane za odpowiednie do tego celu, były te, które ogłosiły bankructwo i miały spore problemy finansowe od kilku lat. Nazwy firm poddanych analizie to: Amber Gold Sp. z o.o., Bielmlek Spółdzielnia Mleczarska, Ursus S.A. oraz Intra Sp. z o.o. Zostały one przebadane w zakresie czasowym obejmującym lata 2016-2021.

Następnie obliczone zostały wskaźniki modeli predykcji upadłości. Posłużyły one do wyliczenia kompletnych wyników w każdym analizowanym roku, zawierających informację, czy dany system dobrze przewidział sytuację spółki. Ze względu na dostępność danych ostatecznie przebadanych zostało łącznie pięć modeli (Hołdy, Gajdki i Stosa, dwa modele Prusaka oraz model „poznański”).

Trzecim etapem było stworzenie pliku z informacjami, na których opierać się miała kreacja modeli Chatu GPT. Zawierał on analizę wad i zalet wybranych do badania klasycznych modeli predykcji upadłości przedsiębiorstw.

Ostatecznym krokiem badawczym była analiza modeli zaproponowanych przez LLM w oparciu o dane wspomnianych wyżej spółek, które zbankrutowały, aby sprawdzić i porównać skuteczność przewidywań.

Dla analizowanych modeli zrealizowany został krótki opis każdego z nich, w którym także oceniono je według następujących kryteriów:

- **dostępność wymaganych danych wejściowych** – tutaj, jeżeli model został wzięty pod uwagę w badaniu, to znaczy, że udało się znaleźć dane w źródłach dostępnych dla studentów Uniwersytetu Ekonomicznego we Wrocławiu. Przykładem, dla którego uzyskanie danych było niemożliwe ze względu na niedostępność danych, jest oprócz innych model Mączyńskiej i Zawadzkiego;
- **skuteczność predykcji** – procentowo wyrażony stosunek liczby uzyskanych obserwacji do wszystkich możliwych do uzyskania wyników dla danego modelu;
- **interpretowalność wyników** – jest to stosunek liczby uzyskanych obserwacji w „szarej strefie” do wszystkich możliwych do uzyskania wyników dla danego modelu;
- **moment, w którym dany model przewidział upadłość przedsiębiorstwa** – w tym kryterium branych pod uwagę jest ostatnich 5 lat działalności podmiotu oraz musi wystąpić minimum jedna prawdziwa obserwacja dla danego modelu.

Poniższa analiza została załączona w pliku „kontekst.txt” „Chatowi GPT” podczas „promptowania”, tj. wydawania polecenia.

Model Hołdy

Model opiera się na liniowej analizie dyskryminacyjnej. Artur Hołda przeanalizował 40 zbankrutowanych firm oraz 40 w dobrej kondycji finansowej (Kisielińska i Waszkowski, 2010, s. 19).

Badanie miało trzy etapy, w których stopniowo eliminowano wskaźniki w celu precyzyjniejszej selekcji. Ostatecznie wyodrębniono pięć kluczowych zmiennych: PWP (podstawowy wskaźnik płynności), SZ (stopa zadłużenia), ZM (zyskowność majątku), WOZ (wskaźnik obrotu zobowiązań) i RM (rotacja majątku). Dyskryminacyjna funkcja liniowa:

$$ZH = 0,605 + 0,681 \times PWP - 0,0196 \times SZ + 0,00969 \times ZM + 0,000672 \times WOZ + 0,157 \times RM, \quad (1)$$

gdzie: PWP = aktywa obrotowe/zobowiązania krótkoterminowe, SZ = zobowiązania ogółem/suma bilansowa, ZM = wynik finansowy netto/średnioroczny majątek ogółem, WOZ = przeciętny stan zobowiązań krótkoterminowych/(koszty działalności operacyjnej – pozostałe koszty operacyjne), RM = przychody ogółem/średnioroczny majątek ogółem.

a) Skuteczność predykcji w analizie autora modelu:

Łącznie – 75%

Bankruci – 50%

Firmy w dobrej kondycji finansowej – 100%.

Skuteczność predykcji w analizie autorów niniejszego rozdziału:

Łącznie – 0%

Bankruci – 0%.

b) Interpretowalność wyników:

ZH < -0,3 – wysoki poziom zagrożenia upadłością

ZH > 0,1 – niewielki poziom zagrożenia upadłością

ZH należy do <-0,3; 0,1> – „szara strefa” (wyniki o wątpliwej jakości).

c) Dostępność danych

Wskaźniki w modelu Hołdy są możliwe do obliczenia za pośrednictwem danych znajdujących się w szczegółowych sprawozdaniach finansowych firm, szczególnie w rachunku zysków i strat oraz bilansie przedsiębiorstwa. Dane te to np. przychody ogółem ze sprzedaży, zobowiązania, suma bilansowa, wynik finansowy netto, poziom aktywów obrotowych, pozostałe koszty operacyjne itp.

d) Moment, w którym model zaczyna przewidywać upadłość

ZH < 0 – ryzyko upadku

ZH > 0 – dobry standing firm.

Model Gajdki i Stosa

Jest to kolejny model podlegający pod wielowymiarową analizę dyskryminacyjną, który pasuje do warunków polskiego rynku. Został przeznaczony do oceny kondycji finansowej spółek notowanych na GPW (Kisielińska i Waszkowski, 2010, s. 20). W modelu zbadanych zostało 17 upadłych przedsiębiorstw, które zostały zestawione z 17 jednostkami nieupadłymi o podobnej działalności.

Dyskryminacyjna funkcja liniowa:

$$Z = -0,0005 \times X_1 + 2,0552 \times X_2 + 1,7260 \times X_3 + 0,1155 \times X_4, \quad (2)$$

gdzie: X_1 = zobowiązania krótkoterminowe/koszt wytworzenia produkcji sprzedanej, X_2 = wynik netto/suma bilansowa, X_3 = wynik brutto/przychody netto ze sprzedaży, X_4 = suma bilansowa/zobowiązania ogółem.

a) Skuteczność predykcji w analizie autora modelu:

Łącznie – 93,8%

Bankruci – 87,5%

Firmy w dobrej kondycji finansowej – 100%.

Skuteczność predykcji w analizie autorów niniejszego rozdziału:

Łącznie: 55% (w zaokrągleniu)

Bankruci: 55% (w zaokrągleniu).

b) Interpretowalność wyników:

$Z < -0,49$ – wysoki poziom zagrożenia upadłością

$Z > 0,49$ – niewielki poziom zagrożenia upadłością

Z należy do <-0,49; 0,49> – „szara strefa” (poziom zagrożenia nieokreślony).

c) Dostępność danych

Podobnie jak w modelu Hołdy dane do obliczenia wskaźników są łatwo dostępne i obejmują zobowiązania, przychody netto czy wynik finansowy. Wyjątek stanowi „koszt wytworzenia produkcji sprzedanej”, trudny do znalezienia w podstawowych sprawozdaniach, ale możliwy do zastąpienia inną, podobną daną, np. „kosztami produkcji”.

d) Moment, w którym model zaczyna przewidywać upadłość

$Z < 0$ – ryzyko upadku firmy

$Z > 0$ – dobry standing firmy.

Modele Prusaka

Błażej Prusak stworzył dwa systemy wczesnego ostrzegania, również oparte na analizie dyskryminacyjnej, z których jeden pozwala na predykcję bankructwa z rocznym wyprzedzeniem, natomiast drugi prognozuje je na 2 lata przed pojawieniem się niepokojącej sytuacji (Wojnar, 2014, s. 224). Postać funkcji obu modeli jest następująca:

$$Z(P1) = -1,5685 + 6,5245 \cdot X1 + 0,148 \cdot X2 + 0,4061 \cdot X3 + 2,1754 \cdot X4, \quad (3)$$

gdzie: $X1$ = zysk operacyjny/suma bilansowa, $X2$ = koszty operacyjne/zobowiązania krótkoterminowe, $X3$ = aktywa obrotowe/zobowiązania krótkoterminowe, $X4$ = zysk operacyjny/przychody ze sprzedaży.

$$Z(P2) = -1,8713 + 1,4383 \cdot X1 + 0,1878 \cdot X2 + 5,0229 \cdot X3, \quad (4)$$

gdzie: $X1$ = (zysk netto + amortyzacja) / zobowiązania ogółem, $X2$ = koszty operacyjne/zobowiązania krótkoterminowe, $X3$ = zysk ze sprzedaży/suma bilansowa.

Badanie na obu modelach obejmowało w próbie uczącej 40 przedsiębiorstw w dobrej kondycji oraz 40 zagrożonych upadłością dobrane parami według rodzaju działalności i branży oraz 39 tak samo dobranych jednostek w próbie testowej.

a) Skuteczność predykcji P1 w analizie autora modelu:

Łącznie – 81,3%

Bankruci – 87,5%

Firmy w dobrej kondycji finansowej – 75%.

Skuteczność predykcji P1 w analizie autorów niniejszego rozdziału:

Łącznie: 70% (w zaokrągleniu)

Bankruci: 70% (w zaokrągleniu).

Skuteczność predykcji P2 w analizie autora modelu:

Łącznie – 50%

Bankruci – 87,5%

Firmy w dobrej kondycji finansowej – 12,5%.

Skuteczność predykcji P2 w analizie autorów niniejszego rozdziału:

Łącznie – 100%

Bankruci – 100%.

b) Interpretowalność wyników P1: $Z < -0,13$ – wysoki poziom zagrożenia upadłością $Z > 0,65$ – niewielki poziom zagrożenia upadłością Z należy do $<-0,13; 0,65>$ – „szara strefa” (poziom zagrożenia nieokreślony).**Interpretowalność wyników P2:** $Z < -0,7$ – wysoki poziom zagrożenia upadłością $Z > 0,2$ – niewielki poziom zagrożenia upadłością.**c) Dostępność danych**

Wszystkie dane, które są potrzebne do obliczenia wskaźników z obu modeli, są stosunkowo łatwo dostępne we wszelkich sprawozdaniach finansowych przedsiębiorstw. Są to np. zobowiązania krótkoterminowe, zysk operacyjny brutto, suma aktywów obrotowych, suma bilansowa, amortyzacja.

d) Moment, w którym model zaczyna przewidywać upadłość dla P1: $Z < -0,13$ – ryzyko upadku firmy $Z > -0,13$ – dobry standing firmy.**Moment, w którym model zaczyna przewidywać upadłość dla P2:** $Z < -0,295$ – ryzyko upadku firmy $Z > -0,295$ – dobry standing firmy.***Model „poznański”***

Model „poznański” autorstwa Hamrola, Czajki i Piechockiego został opracowany na podstawie analizy finansowej. Badania objęły 100 polskich przedsiębiorstw – połowa z nich była w dobrej kondycji finansowej, a wobec pozostałych wszczęto postępowanie upadłościowe lub układowe (Wojnar, 2014, s. 224). Spośród 31 analizowanych wskaźników do modelu wybrano 4 kluczowe, a jego funkcja została przedstawiona następująco:

$$W = -2,368 + 3,562 \cdot X_1 + 1,588 \cdot X_2 + 4,288 \cdot X_3 + 6,719 \cdot X_4 \quad (5),$$

gdzie: X_1 = wynik finansowy netto/majątek całkowity, X_2 = (majątek obrotowy – zapasy) /zobowiązania krótkoterminowe, X_3 = kapitał stały/aktywa, X_4 = wynik finansowy ze sprzedaży/przychody ze sprzedaży.

a) Skuteczność predykcji w analizie autora modelu:

Łącznie – 75%

Bankruci – 50%

Firmy w dobrej kondycji finansowej – 100%.

Skuteczność predykcji w analizie autorów niniejszego rozdziału:

łącznie – 75%

Bankruci – 75%.

b) Interpretowalność wyników:

$Z < 0$ – wysoki poziom zagrożenia upadłością

$Z > 0$ – niewielki poziom zagrożenia upadłością.

c) Dostępność danych

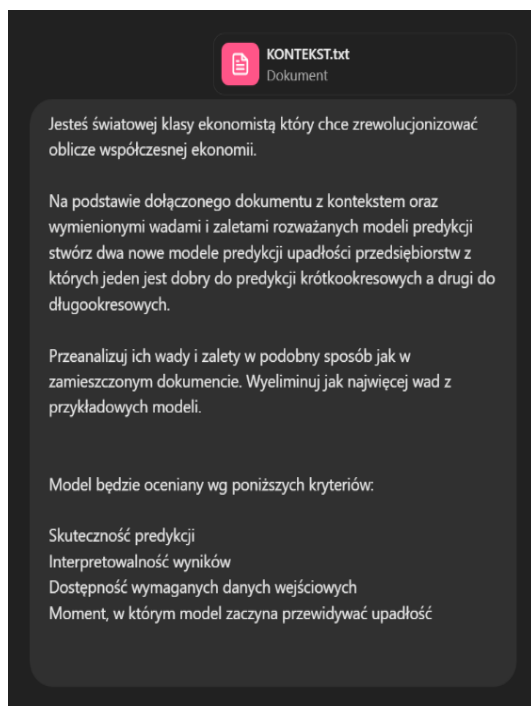
Dane łatwo dostępne na eKRS po wyszukaniu przedsiębiorstwa za pomocą NIP-u oraz pobraniu jego sprawozdań finansowych z konkretnych lat.

d) Moment, w którym model zaczyna przewidywać upadłość

$Z < 0$ – ryzyko upadku firmy

$Z > 0$ – dobry standing firmy.

ChatGPT zaproponował dwa modele: krótkookresowy „KROK” oraz średniookresowy „PERSPEKTYWA”. Pełną odpowiedź chatu przedstawiono na rysunkach 1 i 2.



Rys. 1. Kontekst oraz „prompty” wprowadzone do Chatu GPT potrzebne do wygenerowania nowych modeli

Źródło: opracowanie własne na podstawie „prompta” wpisanego do Chatu GPT.

Oto dwa nowe modele predykcji upadłości przedsiębiorstw:

1. Model KROK (Krótkookresowy Model Ostrzegawczy) – skupia się na predykcji upadłości w perspektywie do 12 miesięcy.
2. Model PERSPEKTYWA – długookresowy model przeznaczony do prognozowania zagrożenia upadłością na 2-5 lat do przodu.

Model KROK – Krótkookresowy Model Ostrzegawczy

Założenia:

- Wykorzystuje dane dostępne w bieżących sprawozdaniach finansowych (np. miesięczne, kwartalne).
- Kładzie nacisk na płynność finansową oraz bieżące zobowiązania.
- Stosuje wskaźniki operacyjne, które szybko reagują na zmiany w firmie.
- Jest oparty na regresji logistycznej, co poprawia interpretowalność wyników.

Funkcja modelu:

$$Z_K = -1,2 + 5,4 \cdot X_1 + 2,8 \cdot X_2 + 1,5 \cdot X_3 + 3,2 \cdot X_4$$

gdzie:

- X_1 – wskaźnik płynności gotówkowej (średni pieniężne / zobowiązania krótkoterminowe).
- X_2 – rotacja należności (przychody ze sprzedaży / należności krótkoterminowe).
- X_3 – wskaźnik rentowności operacyjnej (zysk operacyjny / suma bilansowa).
- X_4 – pokrycie odsetek (EBIT / koszty finansowe).

Interpretacja wyników:

- $Z_K < -0,5$ – wysokie ryzyko upadłości w ciągu najbliższego roku.
- $-0,5 \leq Z_K \leq 0,5$ – sytuacja niepewna (wymagana dalsza analiza).
- $Z_K > 0,5$ – niewielkie ryzyko upadłości.

Zalety:

- ✓ Wysoka skuteczność dla krótkiego horyzontu czasowego (prognozy do 12 miesięcy).
- ✓ Oparty na łatwo dostępnych danych finansowych (np. z rachunku zysków i strat).
- ✓ Wskaźniki szybko reagują na zmiany w działalności operacyjnej.
- ✓ Prosta interpretacja – podział na trzy kategorie ryzyka.

Wady:

- ✗ Nie uwzględnia czynników długoterminowych (np. trendów w zadłużeniu).
- ✗ Może być mniej skuteczny w branżach sezonowych, gdzie wyniki fluktuują.
- ✗ Brak uwzględnienia czynników jakościowych (np. zarządzania, zmian rynkowych).

Model PERSPEKTYWA – Długookresowy Model Prognozy Upadłości

Założenia:

- Analizuje dane z ostatnich 3-5 lat, aby ocenić długoterminowe ryzyko upadłości.
- Uwzględnia zarówno wskaźniki finansowe, jak i czynniki strategiczne.
- Wykorzystuje algorytmy uczenia maszynowego (np. drzewa decyzyjne) do określenia trendów w kondycji przedsiębiorstwa.

Funkcja modelu:

$$Z_P = -2,1 + 4,2 \cdot Y_1 + 3,6 \cdot Y_2 + 2,9 \cdot Y_3 + 1,8 \cdot Y_4 + 2,5 \cdot Y_5$$

gdzie:

- Y_1 – trend zadłużenia (zmiana długu ogółem w % na przestrzeni ostatnich 5 lat).
- Y_2 – stabilność marży operacyjnej (średnie odchylenie standardowe marży EBIT w %).
- Y_3 – dynamika przychodów (średnioroczna zmiana przychodów w %).
- Y_4 – rentowność kapitału własnego (ROE).
- Y_5 – udział zobowiązań długoterminowych w kapitale ogółem.

Interpretacja wyników:

- $Z_P < -1,0$ – wysokie ryzyko upadłości w ciągu 2-5 lat.
- $-1,0 \leq Z_P \leq 1,0$ – umiarkowane ryzyko upadłości (wymagana dodatkowa analiza).
- $Z_P > 1,0$ – niskie ryzyko upadłości.

Zalety:

- ✓ Uwzględnia długoterminowe trendy finansowe.
- ✓ Wykorzystuje analizę szeregów czasowych, co zwiększa dokładność predykcji.
- ✓ Może być stosowany dla dużych przedsiębiorstw i firm z długim okresem działalności.
- ✓ Obejmuje szerszy zakres czynników niż klasyczne modele dyskryminacyjne.

Wady:

- ✗ Wymaga danych historycznych z kilku lat, co może ograniczać jego stosowanie dla nowych firm.
- ✗ Skomplikowana interpretacja wyników – zależność od trendów i zmian rynkowych.
- ✗ Potrzebna regularna aktualizacja modelu w związku z dynamicznymi zmianami rynkowymi.

Rys. 2. Modele wygenerowane przez ChatGPT po wprowadzeniu stosownego „prompta” oraz odpowiednich danych

Źródło: opracowanie własne na podstawie „prompta” wpisanego do Chatu GPT.

Poddane podobnej analizie prezentują się następująco:

Model „KROK” – funkcja zaprezentowana na rysunku 2

a) Skuteczność predykcji w analizie autorów niniejszego rozdziału:

łącznie – 21%
Bankruci – 21%.

b) Interpretowalność wyników:

Została również (razem z funkcją) podana i wyjaśniona na rysunku 2 (znajdującym się na poprzedniej stronie).

c) Dostępność danych

Wszystkie dane były możliwe do znalezienia w sprawozdaniach finansowych (rachunkach zysków i strat, bilansach) spółek wybranych do badania.

d) Moment, w którym model zaczyna przewidywać upadłość

$Z_K < -0,5$ – ryzyko upadku firmy
 $Z_K > 0,5$ – dobry standing firmy.

Model „PERSPEKTYWA” – funkcja zaprezentowana na rysunku 2

a) Skuteczność predykcji w analizie autorów niniejszego rozdziału:

Łącznie – 50%

Bankruci – 50%.

b) Interpretowalność wyników:

Została również (razem z funkcją) podana i wyjaśniona na rysunku 2 (znajdującym się na poprzedniej stronie).

c) Dostępność danych

Dane do obliczenia wskaźników modeli możliwe do odnalezienia. Trudność jednak można odczuć w ich wyliczaniu, co jest czasochłonne oraz wymaga od badacza wiedzy na ich temat, tzn. w jaki sposób je obliczyć i czy podane dane są wystarczające.

d) Moment, w którym model zaczyna przewidywać upadłość

$Z_k < -1$ – ryzyko upadku firmy

$Z_k > 1$ – dobry standing firmy.

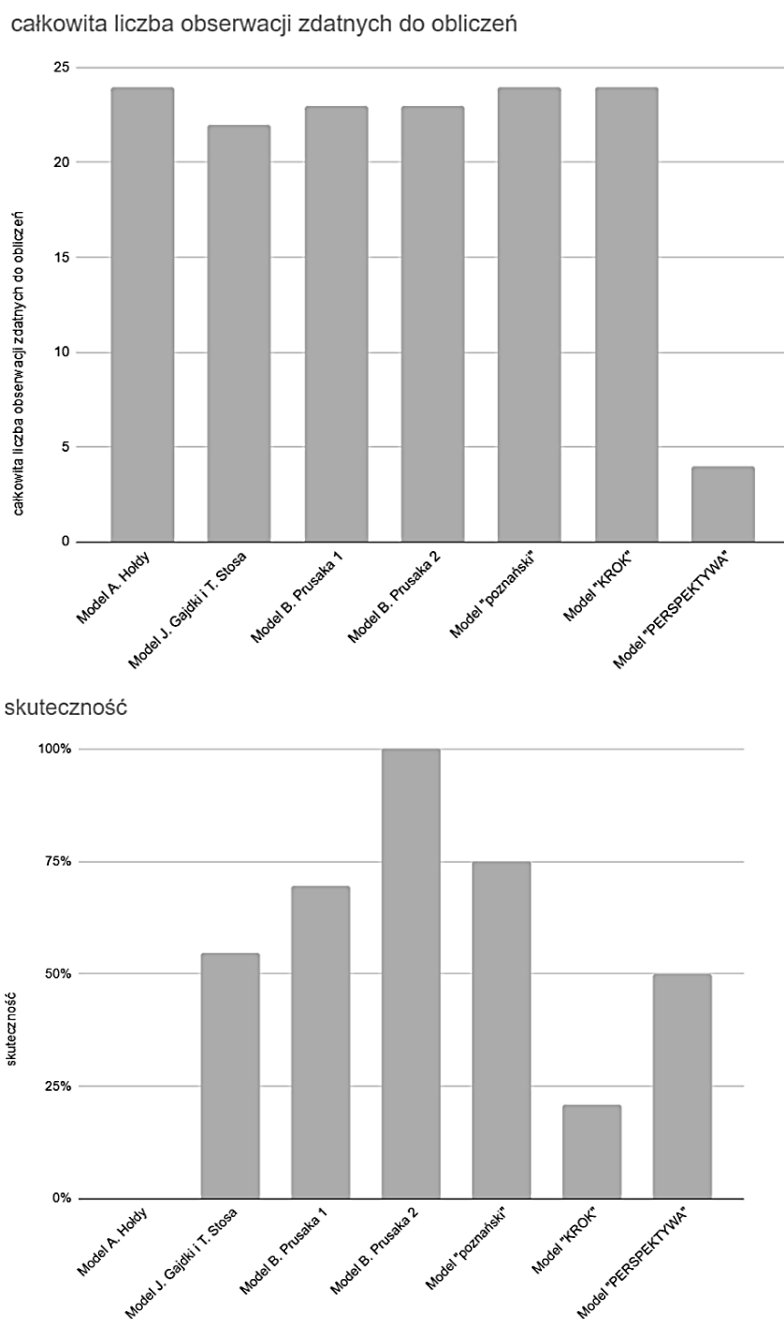
4. Wyniki badań

Dane z bazy zostały przefiltrowane ręcznie pod kątem przydatności do obliczeń w modelach. Były to sprawozdania finansowe upadłych przedsiębiorstw z lat 2016–2021, a także utworzone samodzielnie tabele z wzorami funkcji dyskryminacyjnych wszystkich badanych modeli predykcji upadłości. Następnie zestawiono je w programie Microsoft Excel, co umożliwiło automatyzację obliczeń dla kolejnych lat i firm. Przedstawione poniżej rezultaty zostały uzyskane po wyliczeniu wszystkich wskaźników modeli zgodnie z określonymi pozycjami w rachunkach zysków i strat, a także w bilansach analizowanych upadłych spółek (tj. Amber Gold Sp. z o.o., Bielmlek Spółdzielnia Mleczarska, Ursus S.A. oraz Intra Sp. z o.o.), w każdym badanym roku (niestety w niektórych latach przy określonych modelach dane były niedostępne dla wybranych spółek, stąd liczba obserwacji zdalnych do obliczeń przy nich się różni). Ostatecznie przydzielono klasy wynikom i określono skuteczność modeli, a wyniki przedstawiono w sporządzonej w Excelu tabeli 1.

Tabela 1. Całkowita liczba obserwacji zdalnych do obliczeń, liczba obserwacji prawdziwych oraz skuteczność analizowanych modeli

Nazwa modelu	Całkowita liczba obserwacji zdalnych do obliczeń	Liczba obserwacji prawdziwych	Skuteczność
Model A. Hołdy	24	0	0%
Model J. Gajdki i Stosa	22	12	55%
Model B. Prusaka 1	23	16	70%
Model B. Prusaka 2	23	23	100%
Model „poznański”	24	18	75%
Model „KROK”	24	5	21%
Model „PERSPEKTYWA”	4	2	50%

Źródło: opracowanie własne na podstawie zgromadzonych danych ze sprawozdań finansowych upadłych spółek wybranych do badania.



Rys. 3. Całkowita liczba obserwacji zdalnych do obliczeń oraz skuteczność wybranych modeli predykcji upadłości przedstawione na wykresie

Źródło: opracowanie własne na podstawie danych zgromadzonych ze sprawozdań finansowych upadłych spółek wybranych do badania.

Model „PERSPEKTYWA” jest przeznaczony do przewidywania bankructwa spółki na 3-5 lat przed owym stanem. Z tego też powodu do jego analizy zostało przyjętych ostatnich 5 lat sprawozdań finansowych przedsiębiorstw, z których zostały wyselekcjonowane dane, na podstawie których uzyskano wyniki dla każdej obserwacji dla lat 2017-2021, co dało łączną całkowitą liczbę możliwych obserwacji na poziomie czterech (4 analizowane spółki) przy tym rozwiązaniu. Jak można zaobserwować w tabeli 1, najwyższym poziomem skuteczności spośród analizowanych modeli odznaczył się drugi z modeli Prusaka, natomiast najniższy (zerowy) poziom skuteczności wykazał model Hołdy. Graficzne zobrazowanie całkowitej liczby obserwacji zdalnych do obliczeń oraz skuteczność podanych w tabeli 1 modeli widać na powyższym rysunku 3.

Ograniczona liczba obserwacji przedstawiona na rysunku 3 wpłynęła także na maksymalną liczbę prawdziwych wyników. Analizując procentowy udział prawdziwych obserwacji, można zauważyć, że modele stworzone przez ChatGPT zajęły przedostatnią oraz trzecią od końca pozycję, co zostało pokazane na rysunku 3.

Należy zaznaczyć, że model Hołdy w tym badaniu osiągnął najgorszy wynik – każda z 24 obserwacji była fałszywa. Drugi najniższy wynik spośród analizowanych modeli należy do narzędzia stworzonego przez LLM – modelu „KROK”.

Statystyczna skuteczność dla modeli klasycznych wynosi 60%, zatem oba modele zaproponowane przez ChatGPT znajdują się poniżej średniego wyniku modelu na badanym zestawie danych.

5. Wnioski i rekomendacje

Celem badawczym było porównanie skuteczności polskich modeli dyskryminacyjnych w prognozowaniu upadłości z modelami stworzonymi przez LLM. Badanie miało ocenić, czy nowoczesne modele językowe mogą konkurować z tradycyjnymi metodami, a także ustalić ich umiejętność przewidywania, mierzoną skutecznością predykcji bankructwa spółek.

Wyniki wskazują, że ChatGPT ma potencjał w budowie modeli predykcyjnych, osiągając lepsze wyniki niż klasyczny model Hołdy. Biorąc jednak pod uwagę brak specjalistycznego szkolenia LLM w tymże kierunku, może to wynikać z losowości, co wymaga dalszych badań na większym zbiorze danych. Wówczas będzie można ustalić, czy takie narzędzia mogą znaleźć swoje zastosowanie w świecie finansów i realnie wspomóc menadżerów finansowych w ich analizach. Obecnie drzemie bardzo duży potencjał w tym narzędziu – kontekstowe wzbogacanie o czynniki pozafinansowe lub zachęcenie LLM do wyszukania ich we własnym zakresie oraz dopasowanie modelu do realiów polskiego rynku odpowiednim „promptem” zdecydowanie zwiększyłoby jego przydatność z punktu widzenia zarządzania finansami oraz wyeliminowałoby jedną z wad „klasycznych” modeli predykcji upadłości przedsiębiorstw. Należy także zaznaczyć, że rezultaty osiągnięte przez modele „KROK” i „PERSPEKTYWA” nie należą do najlepszych, ale może to wynikać z niskiego wyspecjalizowania Chatu GPT w tej dziedzinie, braku jego wglądu do bieżących i przeszłych sprawozdań

finansowych spółek (co znacznie zmniejsza zakres przydatnych informacji do wygenerowania odpowiedniego modelu) czy też niemożności uwzględnienia polskich warunków ekonomicznych. Zatem przy obecnym stanie narzędzia nie zaleca się jego stosowania w celach innych niż doświadczalne. Sytuacja może ulec zmianie po wyspecjalizowaniu LLM. Należy jednak pamiętać, że autorzy niniejszego rozdziału odradzają stosowanie modeli predykcji upadłości wygenerowanych przez AI w bieżącym czasie, a nie korzystanie ze wsparcia, jakie oferuje to nowoczesne narzędzie.

W przyszłych badaniach ważne jest także minimalizowanie błędów ludzkich, np. podczas ręcznego przepisywania danych, gdyż wpływają one na rzetelność analizy. Dodatkowo porównanie z większą liczbą klasycznych modeli mogłoby lepiej określić pozycję LLM w prognozowaniu upadłości.

Literatura

- Banach-Kobyra, A. (2017). Prognozowanie upadłości przedsiębiorstw jako narzędzie monitorowania zaistnienia sytuacji kryzysowych. *Przedsiębiorstwo we współczesnej gospodarce – teoria i praktyka*, 3(22), 213-230.
- Kisielińska, J. i Waszkowski, A. (2010). Polskie modele do prognozowania bankructwa przedsiębiorstw i ich weryfikacja. *Zeszyty Naukowe SGGW – Ekonomika i Organizacja Gospodarki Żywnościowej*, (82), 17-31.
- Kocoń, J., Cichecki, I., Kochanek, M., Szydło, D., Baran, J., Bielaniec, J., Gruza, M., Janz, A., Kanclerz, K., Kocoń, A., Koptyra, B., Mieleśzczenko-Kowszewicz, W., Miłkowski, P., Oleksy, M., Piasecki, M., Radliński, Ł., Wojtasik, K., Woźniak, S. i Kazienko, P. (2023, 27 czerwca). *ChatGPT: Jack of All Trades, Master of None*. arXiv Pobrano 28 lutego 2025 z <https://arxiv.org/pdf/2302.10724>
- Piszczek, R. (2009). Zastosowanie modelu logit w modelowaniu upadłości. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, (76), 133-148.
- Tłuczak, A. (2013). Zastosowanie dyskryminacyjnych modeli przewidywania bankructwa do oceny ryzyka upadłości przedsiębiorstw. *Zeszyty Naukowe Wyższej Szkoły Bankowej we Wrocławiu*, 2(34), 423-434.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. i Polosukhin, I. (2017, 12 czerwca). *Attention is All you Need*. NIPS papers. Pobrano 28 lutego 2025 z https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf
- Wojnar, J. (2014). Ocena skuteczności modeli analizy dyskryminacyjnej do prognozowania zagrożenia finansowego spółek giełdowych. *Zeszyty Naukowe Małopolskiej Wyższej Szkoły Ekonomicznej w Tarnowie*, 24(1), 219-231.

The Potential of Artificial Intelligence in Forecasting Corporate Bankruptcy – Comparative Analysis of Discriminatory Bankruptcy Prediction Models of Polish Enterprises

Abstract: Predicting the risk of corporate bankruptcy is a key challenge for managers and financial institutions. Accurate risk assessment enables the implementation of appropriate preventive measures. There are many bankruptcy prediction models based on financial analysis, discriminant analysis, and statistical methods. However, this section of the monograph focuses on comparing traditional Polish

discriminant models (e.g., those by J. Gajdka and T. Stos) with modern solutions proposed by *ChatGPT* – a model commonly referred to as „artificial intelligence” and categorised as an LLM (*Large Language Model*). The comparison is based on the effectiveness of their predictions. The aim of this article is to test these tools in practice, assess their potential, and determine whether they can be applied in financial management. The analysis indicates that AI-driven early warning systems can be more effective than some traditional models and show promise in bankruptcy risk forecasting. The research conducted by the authors of this article focused on selected bankrupt companies and was based on comparative analysis, data and financial statement analysis, critical literature review, evaluation of the effectiveness of various bankruptcy prediction models, and the use of language models.

Keywords: bankruptcy prediction models, ChatGPT, bankrupt companies, discriminant analysis