

Kraków, 14.03. 2017.

Prof. dr hab. Józef Pociecha  
Uniwersytet Ekonomiczny w Krakowie

### **Recenzja**

#### **pracy doktorskiej mgr Damiana Gąski pt. „Przewidywanie bankructwa przedsiębiorstw za pomocą metod uczenia”**

##### 1. Problematyka i struktura pracy

Przedstawiona rozprawa jest dziełem dość obszernym, liczy bowiem 209 stron. Praca oparta jest na literaturze, obejmującej 122 pozycje, w tym 71 to pozycje w języku angielskim. Na tej liście znajdują się podstawowe pozycje literatury światowej z zakresu uczenia maszynowego oraz rozpoznawania obrazów.

Modelowanie i prognozowanie upadłości przedsiębiorstw jest jednym z kluczowych i trudniejszych problemów analiz mikroekonomicznych. Z tego punktu widzenia podjęcie tematu modelowania upadłości przedsiębiorstw jako tematu pracy doktorskiej uważam za bardzo trafne i mające duże znaczenie praktyczne. Spośród wielu metod modelowania i prognozowania upadłości firm w ostatnim czasie zyskują znaczenie statystyczne algorytmy uczenia maszynowego. Wchodzą one w skład obecnie szybko rozwijającej się gałęzi wiedzy jaką są systemy sztucznej inteligencji. Algorytmy uczenia maszynowego pozwalają na adaptacyjne zmiany systemu, tak aby wykonywał on za następnym razem takie samo zadanie lub zadania podobne ale bardziej efektywnie. Wśród metod uczenia maszynowego wyróżnia się uczenie nadzorowane (z nauczycielem), które zakłada obecność ludzkiego nadzoru nad tworzeniem funkcji odwzorowującej wejście systemu na jego wyjście oraz uczenie nienadzorowane (bez nauczyciela), które zakłada brak obecności dokładnego lub nawet przybliżonego wyjścia w danych uczących. Przedmiotem pracy są metody uczenia nadzorowanego, takie jak: naiwny klasyfikator bayesowski, sieci bayesowskie, model regresji logistycznej, metody wektorów nośnych, metoda k-najbliższych sąsiadów, funkcje potencjałowe, klasyfikatory jądrowe, drzewa decyzyjne, lasy losowe i sztuczne sieci neuronowe. Reasumując, wybór tematu pracy doktorskiej oraz metod zastosowanych do rozwiązania postawionych problemów uważam za trafny i mający duże znaczenie praktyczne.

Praca oparta jest na solidnej podstawie literaturowej. Należy podkreślić, że autor swoje wywody opiera na klasycznych pracach z zakresu statystycznego uczenia maszynowego np. takich jak: V. Vapnik (1995), „The Nature of Statistical Learning Theory”, L. Devroye, L.



Gyorfi, G. Lugosi (1996), „Probabilistic Theory of Pattern Recognition”, T. Mitchell (1997), „Machine Learning, C. M. Bishop (2006), „Pattern Recognition and Machine Learning”, T. Hastie, R. Tibshirani, J. Friedman (2009), „The Elements of Statistical Learning: Data Mining, Inference and Prediction”, Prezentowanie rozważań według tych klasycznych w świecie prac uważam za atut ocenianej rozprawy doktorskiej. Podobnie jeśli chodzi o literaturę polską, autor opiera się na klasycznych pracach, jak m.in.: R. Tadeusiewicz (1993), J. Ćwik, J. Koronacki (2008), M. Krzyśko, W. Wołyński, T. Górecki, M. Skorzybut (2008).

Praca składa się z przedmowy, pięciu rozdziałów, zakończenia oraz dodatku. W przedmowie sformułowano (może zbyt skrótowo) cel pracy oraz dwie hipotezy badawcze. Nie wspomniano wyraźnie o zakresie pracy, czyli o rozgraniczeniu pomiędzy problemami i metodami omawianymi w pracy a tymi, wiążącymi się z problematyką bankructwa, które zostały pominięte. Na końcu przedmowy scharakteryzowano (także nieco moim zdaniem zbyt skrótowo) strukturę pracy.

Rozdział pierwszy pracy, zatytułowany „Wprowadzenie”, rzeczywiście wprowadza czytelnika w ekonomiczne zagadnienia związane z bankructwem firm i jego przewidywaniem. Na str. 8 błędnie zinterpretowano pochodzenie słowa bankructwo, „banka rotta” – „złamana ławka”. W dawnych czasach operacji bankierskich dokonywano przy ławkach stojących w miejscach publicznych. Ławki niesolidnych bankierów były przełamywane na znak ich nierzetelności i wyeliminowania z rynku. Pierwotnie nie było to więc łamanie ławek do siedzenia przed domami dłużników, gdyż nie zawsze tam się znajdowały. W p. 1.3 Scharakteryzowano użytkowników metod przewidywania bankructwa. Wśród nich tylko w ostatnim zdaniu wymieniono biegłych rewidentów czy firmy auditingowe. A są to również istotni odbiorcy metod przewidywania bankructwa, gdyż zgodnie z obowiązującymi przepisami, są oni obligatoryjnie zobowiązani aby w swojej ocenie sprawozdania finansowego zawrzeć konkluzję co do zdolności do kontynuowania działalności, co w praktyce oznacza że biegły rewident zaręcza iż do końca następnego okresu sprawozdawczego badana jednostka nie zbankrutuje.

Rozdział drugi poświęcony został przedstawieniu zasad i teorii uczenia pod nadzorem. Ogólnie oceniam go jako dobry opis zasad uczenia nadzorowanego, napisany zgodnie z treścią wspomnianych wcześniej klasycznych prac w literaturze światowej. Prowadzone rozważania zostały właściwie sformalizowane i zilustrowane graficznie. Istotnym punktem tego rozdziału jest p. 2.3 przedstawiający klasyfikację metod uczenia nadzorowanego. Jego syntezą jest opracowany przez autora rys. 2.3 przedstawiający drzewo klasyfikacji metod uczenia pod nadzorem.



Najobszerniejszym a zarazem kluczowym rozdziałem pracy jest rozdział trzeci, zawierający opis metod uczenia pod nadzorem. Opis metod rozpoczęto od naiwnego klasyfikatora bayesowskiego. Podejście to jest na ogół pomijane w zastosowaniach metod uczenia statystycznego do przewidywania bankructwa, a może ono stanowić rodzaj punktu odniesienia rezultatów klasyfikacji, gdyż przyjmuje założenie o warunkowej niezależności rozpatrywanych zmiennych. Drugą z zaprezentowanych metod uczenia są sieci bayesowskie, będące graficznym modelem probabilistycznym odwzorowującym strukturę stochastycznych zależności pomiędzy zmiennymi losowymi w postaci grafu. Autor podkreśla, że najistotniejszym i najtrudniejszym problemem uczenia grafu jest uczenie jego struktury. Przedstawia w tym względzie trzy grupy metod uczenia struktury. Pierwsza z nich to metody uczenia struktury oparte na ograniczeniach, druga to algorytmy wartościujące a trzecia to podejście hybrydowe łącznie cechy dwóch pierwszych podejść.

Ważną metodą uczenia maszynowego jest wykorzystanie w tym celu funkcji regresji logistycznej. Tutaj nie podkreślono w wystarczającym stopniu, że modele logitowe stanowią uniwersalne narzędzie klasyfikacji, niekoniecznie związane z algorytmami uczenia. Można było rozszerzyć ten opis przez sprecyzowanie ilorazu szans, stąd wyprowadzenie logitu i zwrócenie uwagi na interpretację modelu logitowego. W wyczerpujący sposób zaprezentowano metodę wektorów nośnych (SVM) jako ważnego i dość uniwersalnego klasyfikatora. Kolejną metodą zaprezentowaną w tym rozdziale jest metoda k-najbliższych sąsiadów.

Szeroką klasą narzędzi uczenia nadzorowanego, są funkcje potencjałowe. Autor klarownie przedstawił ideę, formalizację i prezentację graficzną funkcji potencjałowych. W punkcie tym zawarł również własną propozycję modyfikacji klasyfikatora wykorzystującego funkcje jądrowe w postaci uwzględniania wagi wpływu elementu z próby uczącej nie tylko jego odległością od nowej obserwacji ale również od wektora centralnego klasy z której pochodzi. Rozwinięciem klasycznych klasyfikatorów jądrowych są klasyfikatory jądrowe maksymalnego marginesu, przedstawione w kolejnym punkcie tego rozdziału.

Następną grupą metod uczenia nadzorowanego są drzewa decyzyjne oraz lasy losowe. Metoda drzew decyzyjnych została w pracy przedstawiona w wariacie CART. Za względu na charakter badanej zmiennej (bankrut, niebankrut) drzewa te mają charakter drzew klasyfikacyjnych. Pewnym uogólnieniem metody drzew decyzyjnych są lasy losowe, wykorzystujące bootstrapowe replikacje próby uczącej do budowy całego zbioru drzew. Należy podkreślić, że dość trudna problematyka drzew decyzyjnych i lasów losowych została tutaj przedstawiona w sposób precyzyjny oraz pogładowy.



Kolejną grupą metod uczenia maszynowego są sztuczne sieci neuronowe. Również i tutaj autor szeroko przedstawił ideę sieci neuronowych, ich klasyfikację, architekturę, metody aktywacji, algorytm uczenia sieci oraz wykorzystanie do klasyfikacji obiektów. Przedstawiono zasady oraz algorytm wstecznej propagacji błędu. Na końcu tego punktu autor zestawiał zalety i wady sieci neuronowych jako narzędzi prognozowania bankructwa.

Ostatnim punktem rozdziału prezentującego wachlarz metod uczenia nadzorowanego jest przedstawienie metod łączenia klasyfikatorów w model zagregowany, czyli podejścia wielomodelowego. Podstawowymi metodami agregacji są metody „bagging” i „boosting”. Przedstawiono kolejno: architekturę agregacji, regułę głosowania, algorytm metody bagging, schemat graficzny tej metody. Analogicznie dla metody boosting przedstawiono architekturę szeregową oraz algorytm AdaBoost.

Reasumując ocenę treści rozdziału trzeciego, jako podstawowego rozdziału teoretycznego stwierdzam, że napisany on został ze znanstwem poruszanych zagadnień, w sposób dość precyzyjny, odpowiednio sformalizowany a jednocześnie pogładowy. Uważam że uważny czytelnik tego rozdziału nie będzie miał problemów ze zrozumieniem opisywanych tam metod uczenia nadzorowanego.

W rozdziale czwartym przedstawiono sposoby porównywania efektywności zaprezentowanych metod prognozowania bankructwa. Zaprezentowano w nim procedurę podstawień krzyżowych, za pomocą której można wyznaczyć miary ACC oraz MCC. Pewną alternatywą w stosunku do podstawień krzyżowych jest metoda bootstrap. Tutaj wyznaczamy miary FPR oraz TPR, których zależność można przedstawić graficznie w postaci krzywej ROC oraz pola AUC. Oprócz wymienionych krzywych przedstawiono także wskaźnik Briera. Miary te są adekwatnymi współczesnymi narzędziami oceny jakości procedur klasyfikacyjnych.

W p. 4.2 przedstawiono możliwości zastosowania testów permutacyjnych do oceny jakości rozpatrywanych klasyfikatorów. Zaprezentowano w nim, na podstawie literatury zagranicznej, ideę, zasady oraz algorytm testu permutacyjnego. Tutaj zwracam uwagę, że chociaż prezentacja wywodów nie budzi zastrzeżeń, to pominięto polski dorobek w tym zakresie, a szczególnie książkę G. Kończaka „ Testy permutacyjne. Teoria i zastosowania”, Katowice 2016 oraz jego wcześniejsze publikacje z tego zakresu.

W p. 4.3 autor słusznie zauważył, że efektywność metod uczenia z reguły zależy od wstępnego przetworzenia danych i zaproponował wybór cech wykorzystujący zestawienie metody bootstrap oraz procedury wstecznej eliminacji. Alternatywą jest wybór cech o



największej sile dyskryminacyjnej. Do wstępnego przetwarzania danych proponuje standaryzację lub metodę głównych składowych.

Rozdział piąty ma charakter empiryczny a jego ocenę sformułuję w następnym punkcie. W zakończeniu pracy autor przedstawił syntezę otrzymanych wyników badań empirycznych oraz na końcu zarysował dalsze możliwe kierunki badań. Przedstawioną syntezę uważam za trafną, aczkolwiek wolałbym aby była ona obszerniejsza i aby w szerszym stopniu odnosiła się do poszczególnych metod, prezentowanych w rozdziale trzecim.

## 2. Ocena empirycznej warstwy pracy

Oceniając rozdział piąty pracy, zawierający relację z przeprowadzonych szerokich badań empirycznych, chciałbym zwrócić uwagę na problemy jakie stwarza badana rzeczywistość ekonomiczna. Część teoretyczna pracy została przedstawiona w sposób dość klarowny, elegancki i usystematyzowany. Przechodząc do analiz empirycznych badacz ciągle musi szukać kompromisów pomiędzy wymogami teorii a badaną rzeczywistością. Jest to trudność a jednocześnie urok badań empirycznych. Autor pracy miał z tym ciągle do czynienia i należy go pochwalić, że przeprowadził tak obszerne badania empiryczne a co do konkretnych rozstrzygnięć nie można mieć większych zastrzeżeń.

Rozpoczynając od uwag szczegółowych, wskazuję że w p. 5.1 niepotrzebnie dosłownie powtórzono hipotezy badawcze przedstawione w przedmowie na str. 3. Również zwracam uwagę, że tradycyjny układ grup wskaźników finansowych jest następujący: wskaźniki płynności, zadłużenia, efektywności działania, rentowności, wartości rynkowej (Sierpińska, Jachna, 1997 lub późniejsze wydania, D. Wędzki, 2009). Wskaźniki płynności znajdują się na pierwszym miejscu, gdyż główną przyczyną upadłości firm jest trwała utrata płynności. Zdarza się że firma rentowna upada, gdyż utraciła płynność a firma nierentowna może przez długi czas utrzymywać płynność.

Rys. 5.1 wskazuje jak trudno jest zgromadzić dane dotyczące pełnego, przyjętego w pracy zakresu wskaźników. Podzielam pogląd autora że to mogła być przyczyna braku potwierdzenia drugiej hipotezy badawczej. W pracy wykonano solidną analizę przyjętych zmiennych przez analizę rozkładów cech w grupach spółek zdrowych i bankrutów, zbadano indywidualną siłę dyskryminacyjną cech oraz dokonano prezentacji graficznej w postaci krzywych ROC i pól AUC.

Tabela 5.2 pokazuje jak szeroki był zakres przeprowadzonych badań empirycznych dla poszczególnych klasyfikatorów i ich wariantów. To uzmysławia czytelnikowi jak pracochłonne były przeprowadzone badania empiryczne. Z drugiej strony powoduje to



pewien natłok i chaos przekazanych informacji. Szkoda że nie pomyślano nad jakąś edytorską formą czytelnej syntetycznej prezentacji stopnia przydatności prezentowanych metod uczenia maszynowego do prognozowania bankructwa firm giełdowych.

Skromniej wypadły rezultaty badań empirycznych obejmujących wszystkie pięć grup wskaźników finansowych, czyli po dołączeniu do pierwotnego zbioru cech, wskaźników wartości rynkowej firm. Tutaj wyniki nie są jednoznaczne. W efekcie autor stwierdza że wyniki te nie dają jednak podstaw, by hipotezę o przydatności wskaźników rynku finansowego odrzucić.

Reasumując swoją ocenę części empirycznej pracy podkreślam, że autor podjął się szeroko zakrojonych badań empirycznych, które wykonał z należytą starannością. Otrzymane wyniki, jako że dotyczą skomplikowanej rzeczywistości gospodarczej, nie chcą układać się zgodnie z teoretycznymi regułami wyrażonymi wcześniej w opisie zastosowanych metod. Ale jak to wyraziłem wcześniej jest to urok badań empirycznych. Przedstawioną część empiryczną uważam za silny atut tej pracy.

Z uznaniem należy też odnotować, że obliczenia zostały wykonane przez autora w języku R częściowo na podstawie gotowych funkcji zgromadzonych w odpowiednich pakietach statystycznych a częściowo samodzielnie stworzonych przez autora. Ich opis został zawarty w Dodatku A. Biblioteka funkcji w języku R została dołączona do pracy w postaci pliku na płycie CD. Zgromadzenie i częściowe stworzenie takiej biblioteki jest kolejnym istotnym dorobkiem naukowym autora.

### 3. Uwagi, opinie i pytania

Przestudiowanie tej niewątpliwie ciekawej pracy skłania do sformułowania pewnych uwag i opinii natury bardziej ogólnej, które w pewnej części wynikają z subiektywnego odbioru pracy przez recenzenta.

1. Kwestia tytułu pracy. Tytuł pracy „Przewidywanie bankructwa przedsiębiorstw za pomocą metod uczenia” uważam za zbyt skrótowy. Na końcu, moim zdaniem powinien zostać dodany przymiotnik określający charakter tego uczenia np. „nadzorowanego”, gdyż dla osób nie zajmujących się metodami uczenia maszynowego termin „uczenie” rozumiany jest tradycyjnie, bez związku z prezentowanymi procedurami.
2. We wprowadzeniu oraz w rozdziale trzecim nadmieniono, że model logitowy stanowi uznane narzędzie klasyfikacji i prognozowania bankructwa. W p. 3.3 przedstawiono model regresji logistycznej jako jedną z procedur uczenia nadzorowanego. Brakło mi



szerszej prezentacji we wprowadzeniu do modelu logitowego podkreślenia, że jest to ogólne narzędzie modelowania i prognozowania wielu zjawisk społeczno-ekonomicznych.

3. W p. 1.4 dokonano przeglądu literatury w języku polskim, dotyczącej prognozowania bankructwa. Moim zdaniem jest on zbyt skrótowy i nie przedstawia autorskiej oceny znaczenia powoływanych prac. Pominięto w nim także znaczące prace, takie jak np. prace A. Hołdy (publikowane po 2001 roku) lub prace D. Wędzkiego.
4. W części empirycznej pracy podawano wyłącznie sposoby liczenia poszczególnych wskaźników ( $X_i$ ) a wskaźniki te w analizie finansowej posiadają swoje ustalone nazwy np. wskaźnik  $X_2$  to rentowność majątku (ROA) a  $X_5$  to rentowność kapitału własnego (ROE) itd. Szkoda że oprócz definicji poszczególnych wskaźników nie podano ich nazw ogólnych, przyjętych w analizie finansowej.
5. W analizie zbioru firm w przestrzeni zarówno wskaźników wynikających z analizy sprawozdań finansowych jak i wartości rynkowej firm kluczową rolę odgrywa ich skorelowanie oraz opóźnienie w czasie. Wskaźniki finansowe są liczone na podstawie rocznych sprawozdań finansowych z przeszłości. Inwestorzy na rynku kapitałowym określając cenę rynkową firmy kierują się zarówno informacjami o jej kondycji finansowej w przeszłości, wyrażonej przez odpowiednie wskaźniki finansowe jak i bieżącą informacją mówiącą o nastrojach panujących na rynku finansowym. Powstaje zasadniczy problem jak te dwa rodzaje informacji połączyć ze sobą. W związku z tym chciałbym zapytać o ewentualne sugestie dotyczące możliwości rozwiązania tego problemu na gruncie metod uczenia maszynowego.

Przedstawione powyżej uwagi i opinie uważam za otwarte a stawiam je w intencji podjęcia dyskusji naukowej z osobą, która wykonując obszerne studia literaturowe oraz prowadząc badania empiryczne zdobyła pewne doświadczenie w tym zakresie.

#### 4. Konkluzja

Reasumując ocenę przedstawionej rozprawy doktorskiej stwierdzam że:

- autor podjął bardzo aktualny i trudny temat prognozowania upadłości przedsiębiorstw za pomocą metod uczenia nadzorowanego, na przykładzie spółek giełdowych notowanych na GPW w latach 2008-2016,
- we sposób systematyczny i odpowiednio sformalizowany zaprezentował współczesne metody uczenia nadzorowanego,
- dogłębnie przedstawił metody oceny skuteczności predykcyjnej metod uczenia

nadzorowanego,

- przeprowadził obszerne badania empiryczne dotyczące analizy porównawczej, przedstawionych w części teoretycznej, procedur uczenia nadzorowanego,
- zebrał i częściowo samodzielnie przygotował funkcje języka R do zrealizowania obliczeń empirycznych w pracy.

Tym samym w pełni zrealizował, postawiony w przedmowie, cel pracy.

Konkludując uważam, że przedłożona rozprawa doktorska w pełni odpowiada wymogom stawianym pracom doktorskim, sformułowanym w Ustawie z dnia 14 marca 2003 roku o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki. Wnoszę przeto o dopuszczenie mgr Damiana Gąski do dalszego postępowania związanego z przewodem doktorskim.

Jednocześnie biorąc pod uwagę wysoki poziom naukowy prezentacji części teoretycznej pracy, obszerny i pracochłonny zakres przeprowadzonych badań empirycznych oraz przygotowanie biblioteki funkcji w języku R uważam rozprawę doktorską mgr D. Gąski za wyróżniającą i zasługującą uhonorowaniem odpowiednią nagrodą.



Józef Pocięcha