



**AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA**

**WYDZIAŁ ELEKTROTECHNIKI, AUTOMATYKI, INFORMATYKI
I INŻYNIERII BIOMEDYCZNEJ
KATEDRA AUTOMATYKI I INŻYNIERII BIOMEDYCZNEJ**

AUTOREFERAT ROZPRAWY DOKTORSKIEJ

**BADANIA NAD METODOLOGIĄ PROGNOZOWANIA STANU
MAŁYCH I ŚREDNICH PRZEDSIĘBIORSTW**

MGR INŻ. ANDRZEJ BURDA

Promotor:

Prof. dr hab. inż. Zdzisław S. Hippe

Wyższa Szkoła Informatyki i Zarządzania
w Rzeszowie

Kraków 2013

STRESZCZENIE

Zasadniczym celem rozprawy było zaprojektowanie, zbudowanie i upowszechnienie przyjaznego narzędzia informatycznego, realizującego opracowaną wcześniej, oryginalną metodykę prognozowania stanu **Małych i Średnich Przedsiębiorstw (MSP)**.

Osiągnięcie postawionego celu badawczego zrealizowano kolejno poprzez: przedstawienie informacji na temat specyfiki **MSP** oraz ich opisu w logice atrybutalnej; opis stanu wiedzy na temat prognozowania bankructwa przedsiębiorstw na podstawie doniesień literaturowych; propozycję oryginalnej metodyki prognozowania stanu **MSP** w warunkach niepewności, opartej na wykorzystaniu zespołowej pracy równoległe połączonych heterogenicznych klasyfikatorów, głównie sztucznych sieci neuronowych.

Przeprowadzone badania umożliwiły dokonanie spostrzeżenia, że jakość prognoz dla przypadków kategorii *bankrut* jest wyraźnie gorsza niż dla kategorii *niebankrut*. Zastosowanie autorskiej koncepcji *walidacji kolejkowej* pozwoliło na ustabilizowanie wyników predykcji w dłuższym okresie czasu, a zaproponowanie własnego sposobu wyznaczania wskaźnika jakości predykcji zwiększyło obiektywizację analiz porównawczych modeli, w których występowały niesklasyfikowane przypadki z pozostałymi modelami.

Skuteczność zaproponowanej metodyki oraz zbudowanego według tej koncepcji systemu informatycznego **ProfileSEEKER**, potwierdzono szeregiem eksperymentów obliczeniowych. System ten udostępniono potencjalnym użytkownikom w sieci Internet.

WSTĘP. CEL I ZAKRES PRACY

Wiadomo, że **MSP** stanowią ważny element gospodarki wszystkich krajów Unii Europejskiej, zaś szczególne znaczenie mają w obszarach słabo zurbanizowanych,

na terenie ubogich regionów – u nas w województwie lubelskim i podkarpackim. Przeprowadzona analiza dostępnej literatury na temat prognozowania stanu **MSP**, jednoznacznie wskazuje na brak dedykowanych rozwiązań (algorytmów, systemów informatycznych), przeznaczonych do oceny stanu (*przetrwanie* lub *upadek*) tego typu firm. Jest to przypuszczalnie spowodowane ograniczonym publicznym dostępem do danych o **MSP**.

Można z dużym prawdopodobieństwem założyć, że zostały one skażone działaniami tzw. *kreatywnej księgowości*, opisanymi zwięźle w Rozdziale 2 rozprawy. Rozsądna metoda uporania się z zagadnieniem możliwie wiarygodnej oceny stanu małego i średniego (a właściwie – *rodzinnego*) przedsiębiorstwa, przypuszczalnie może polegać na zastosowaniu najnowszych mechanizmów uczenia maszynowego, wykorzystujących przede wszystkim modele uczenia wygenerowane sztucznymi sieciami neuronowymi, uzupełnione innymi jeszcze metodami ekstrakcji informacji i wiedzy z danych.

Z tego względu, zasadniczym celem badań jest podjęcie próby opracowania metodyki prognozowania stanu **MSP**, a na tej podstawie – zaprojektowanie, zbudowanie i upowszechnienie przyjaznego narzędzia informatycznego, realizującego wspomniane zadanie. Jednak już pierwsze wyniki eksperymentów wykazały, że wiąże się to z przetwarzaniem danych w warunkach niepewności, spowodowanej *kreatywną księgowością*. Uściślając zatem cel badań można powiedzieć, że dotyczy on opracowania nowego algorytmu prognozowania zagrożenia **MSP** bankructwem, stanowiącego podstawę wystarczająco wiarygodnej metodyki oceny stanu (*przetrwanie* lub *upadek*) wymienionych przedsiębiorstw.

Przeprowadzone w ramach niniejszej dysertacji badania, realizowano z zachowaniem zasady *najpierw doświadczenie, a później wniosek*, wynikającej z doświadczeń nauk przyrodniczych, stanowiących obecnie najważniejszy poligon badawczy nowoczesnej informatyki, inżynierii wiedzy, sztucznej inteligencji i systemów informacyjnych.

W Rozdziale 2 pracy „**MSP**- zarys wiadomości. Dane prawdziwe i dane fałszywe” przedstawiono podstawowe informacje o małych i średnich przedsiębiorstwach ze szczególnym uwzględnieniem specyfiki ich gospodarczych właściwości i zachowań w systemie gospodarczym.

Opis zastosowanych algorytmów uczenia maszynowego oraz autorskiej metody walidacji uzyskanych wyników badań, wynikającej ze specyfiki analizowanych baz informacyjnych, przedstawiono w Rozdział 3 zatytułowanym „*Koncepcja realizacji badań. Stosowane metody, algorytmy oraz sposoby oceny modeli uczenia – walidacja kolejkowa*”.

W kolejnym, czwartym rozdziale pracy „*Walidacja niepewnych danych **MSP** przy pomocy indukcyjnych modeli uczenia maszynowego. Modele sztucznych sieci neuronowych*” przedstawiono wyniki badań nad metodyką prognozowania stanu małych i średnich przedsiębiorstw z użyciem wybranych i wcześniej opisanych metod uczenia sztucznych sieci neuronowych.

Uzyskanie jeszcze bardziej wiarygodnych podstaw budowy planowanego systemu doradczego narzuciło potrzebę podjęcia dalszych badań nad analizą zgromadzonych przypadków innymi – od sztucznych sieci neuronowych – metodami uczenia maszynowego. Sięgnięto więc po metody indukcyjnego generowania reguł składniowych. Wyniki tych badań opisano w Rozdziale 5 zatytułowanym „*Walidacja niepewnych danych **MSP** przy pomocy indukcyjnych modeli uczenia maszynowego. modele uczenia w postaci zbioru reguł decyzji*”.

Zaobserwowane w trakcie tej części badań duże zróżnicowanie jakości klasyfikatorów, nasunęło spostrzeżenie o celowości budowy systemu doradczego wspartego koncepcjami modeli zespołowych. W Rozdziale 6 – „*Walidacja niepewnych danych **MSP** przy pomocy zespołu klasyfikatorów*” – przedstawiono wyniki badań zespołów opartych na algorytmie ważonej większości przy zastosowaniu kilku wariantów tego rozwiązania.

Rozdział 7 zatytułowany „*Metodologia prognozowania stanu małych i średnich przedsiębiorstw*” stanowi niejako podsumowanie rozważań zawartych w poprzednich rozdziałach i zakończony jest sformułowaniem nowej metodyki poszukiwania optymalnego systemu predykcji stanu **MSP**.

W Rozdziale 8 „*Implementacja opracowanej metodologii prognozowania stanu MSP. Możliwości i ograniczenia*” opisano implementacyjne aspekty realizacji systemu doradczego zbudowanego w oparciu o zaproponowaną w Rozdziale 7 metodykę.

Ostatni rozdział zatytułowany „*Podsumowanie wyników badań. Wnioski*” zawiera, zgodnie z przyjętą koncepcją, tezy rozprawy oraz wnioski z przeprowadzonych badań.

Do pracy dołączono 5 załączników, gdzie w załączniku: A – wyszczególniono stosowany sprzęt informatyczny i oprogramowanie, w B – opisano badane bazy informacyjne, w C – zdefiniowano własny wskaźnik jakości predykcji oraz przedstawiono jego interpretację, w D – przedstawiono wyniki rozeznania literaturowego na temat algorytmów do prognozowania stanu przedsiębiorstw, a Załącznik E – stanowi zwieszły przewodnik do plików zapisanych na załączonym do rozprawy nośniku DVD.

PODSUMOWANIE

Badania opisane w niniejszej rozprawie były prowadzone z zachowaniem zasady ***najpierw doświadczenie, a później wnioski***. Główny zamysł badań, dopiero po ich zrealizowaniu przybrał formę udowodnionych tez:

1. *Prowadzone badania umożliwiły uzyskanie wglądu w mechanizm budowy sztucznych sieci neuronowych o optymalnej architekturze dostosowanej do objaśniania właściwości opisujących **MSP**,*

2. *Prowadzone badania umożliwiły uzyskanie nowego spojrzenia na sposób uwarunkowania stanu ekonomiczno-finansowego **MSP**,*
3. *Pozytywne wyniki badań, wzmocnione eksperymentami związanymi z optymalizacją generowanych równocześnie reguł składniowych (lub reguł przekonań, w przypadku analizy sprzecznych danych), stały się podstawą utworzenia narzędzia informatycznego do prognozowania stanów **MSP** oraz udostępnienia go przedsiębiorcom w sieci Internet.*

Problem wglądu w mechanizm budowy sztucznych sieci neuronowych o optymalnej architekturze, dostosowanej do objaśniania właściwości opisujących **MSP**, (*Teza 1-sza*), został szczegółowo przedyskutowany w akapicie zatytułowanym *Omówienie uzyskanych wyników* w **Rozdziale 4** opisującym przeprowadzone eksperymenty. Rekapitulując opisane tam spostrzeżenia można wyrazić przekonanie, że uzyskano informacje co do optymalnej liczby neuronów w warstwach ukrytych we wszystkich stosowanych sztucznych sieciach neuronowych.

Uzyskane wyniki badań, potwierdzają wniosek, że predykcja kategorii ***nie-bankrut***, w odniesieniu do wszystkich analizowanych danych, jest znacznie bardziej dokładna w porównaniu do możliwości dokonania predykcji kategorii ***bankrut***. Błąd predykcji kategorii ***nie-bankrut***, ustalony przy pomocy wszystkich stosowanych rodzajów sztucznych sieci neuronowych, wynosi w przybliżeniu tylko 10%, co w odniesieniu do typowych standardów uczenia maszynowego można uznać za wynik w zupełności zadowalający. Jednocześnie można stwierdzić, że przeprowadzone badania faktycznie umożliwiły uzyskanie nowego spojrzenia, (*Teza 2-ga*), na sposób uwarunkowania stanu **MSP** i na zdolność tego typu przedsiębiorstw do kontynuowania działalności gospodarczej. Ma to istotne znaczenie gospodarcze, z uwagi na bezpośrednią zależność między upadłością **MSP**, a możliwością zaistnienia szybkiego, niekontrolowanego wzrostu bezrobocia. To nowe spojrzenie może zatem przyczynić się do spełnienia ważnej funkcji społecznej.

Uzyskanie wiarygodnych podstaw budowy planowanego systemu doradczego narzuciło konieczność podjęcia badań nad analizą zgromadzonych przypadków innymi – od sztucznych sieci neuronowych – metodami uczenia maszynowego. Sensownym kierunkiem uzupełniających eksperymentów, wydała się próba zastosowania metod indukcyjnego generowania reguł składniowych, i zastosowania ich do analizy posiadanych danych o **MSP**. Modele uczenia uzyskane w ten sposób są bardziej zrozumiałe, niejako samo-objaśniają się. Mogą być i są niezbędnym etapem zastosowania koncepcji nowoczesnego współdziałania człowieka z komputerem, poprzez wykorzystanie możliwości wzajemnej interakcji, ukrytych w systemach doradczych realizujących filozofię **WMA** (weighted majority algorithm). Wyniki tych badań potwierdziły wnioski, sformułowane na podstawie zastosowań klasyfikatorów neuronowych. Prognozowanie konceptu ***nie-bankrut***, niezależnie od zastosowanej metody generowania reguł decyzji, okazało się być bardzo wiarygodne. Dla wielu modeli, w szczególności dla reguł składniowych powstałych w wyniku konwersji sieci przekonań Bayes’a, wskaźnik jakości predykcji przekracza wartość 0.90.

Spostrzeżenia na temat zróżnicowania jakości działania badanych klasyfikatorów zespołowych, stały się podstawą do nadania im formy algorytmów i zaimplementowania ich w module prognozującym narzędzia informatycznego (**ProfileSEEKER**). W ten sposób udowodnione zostało pierwotne założenie, (*Teza 3-cia*), że przeprowadzone badania mogą być podstawą utworzenia narzędzia informatycznego do wiarygodnego prognozowania stanów **MSP** oraz udostępnienia go przedsiębiorcom w sieci Internet.

W uzupełnieniu dyskusji podsumowania wyników przeprowadzonych badań, warto zaznaczyć, że wszystkie wygenerowane modele uczenia, pochodzące zarówno z klasyfikatorów autonomicznych, jak i zespołowych, przypuszczalnie dzięki zastosowaniu specyficznej metody walidacji, uzyskały właściwość stabilizacji wyników predykcji w badanym przedziale czasowym. Zdaje się to wskazywać kierunki dalszych badań nad rozwojem koncepcji *walidacji kolejkowej*.

Reasumując – można wyrazić pogląd, że badania realizowane w ramach niniejszej rozprawy umożliwiły opracowanie nowej metodyki oceny zagrożenia **MSP** bankructwem, która – od sformułowania koncepcji i wymagań funkcjonalnych, poprzez rozwiązania strukturalne i aplikacyjne – doprowadziła do zaprojektowania oraz implementacji przyjaznego narzędzia informatycznego **ProfileSEEKER**.

Przeprowadzone badania potwierdzają tym samym spełnienie ogólnej tezy, że jest możliwym opracowanie wystarczająco wiarygodnej metodyki oceny stanu (*prze-trwanie* lub *upadek*) wymienionych przedsiębiorstw, prowadzonej w warunkach niepewności, spowodowanej kreatywną, bądź agresywną księgowością.

WYKAZ ORYGINALNYCH PRAC NAUKOWYCH Z ZAKRESU BADAŃ NAD METODYKĄ PROGNOZOWANIA STANU MAŁYCH I ŚREDNICH PRZEDSIĘBIORSTW, OPUBLIKOWANYCH W WYNIKU REALIZACJI ROZPRAWY

1. A. Burda: *Neuronowy system ostrzegania przedsiębiorstwa przed bankructwem - wybór optymalnej architektury*, **Zamojskie Studia i Materiały** 19(2006)147-156.
2. B. Kuczmowska, A. Burda, Z.S. Hippe: *Prediction of Economic Situation of Small and Medium Enterprises Using Bayesian Network*, W: Kurzynski, M.; Puchala, E.; Wozniak, M.; Zolnierek, A. (Eds.) *Computer Recognition Systems 2*, Springer, Berlin / Heidelberg 2007, str. 802-807.
3. A. Burda, B. Kuczmowska, Z.S. Hippe: *Ensembles of Artificial Neural Networks for Predicting Economic Situation of Small and Medium Enterprises*, W: Kurzynski, M.; Puchala, E.; Wozniak, M.; Zolnierek, A. (Eds.) *Computer Recognition Systems 2*, Springer, Berlin / Heidelberg 2007, str. 808-815.
4. A. Burda, Z.S. Hippe: *Methodology of Predicting the Future Condition of Small and Medium Enterprises*, W: Tadeusiewicz, R.; Ligęza, A.; Mitkowski, W.; Szymkat, M. (Eds.) *7th Conference Computer Methods and Systems*, Oprogramowanie Naukowo Techniczne, Kraków 2009:61-64.
5. A. Burda, Z.S. Hippe: *Uncertain data modeling: The case of small and medium enterprises*, W: Pardela T., Wilamowski B. (Eds.) *Proc. 3rd International Conference on Human System Interaction (HSI'2010)*, Rzeszów 13-15.05.2010, str. 76-80.
6. A. Burda, Z.S. Hippe: *From the research on modeling of uncertain data: The case of small and medium enterprises*, W: Hippe Z.S., Kulikowski J.L., Mroczek T. (Eds.) *Human-Computer Systems Interaction: Backgrounds and Application 2, Part I (Advances in Intelligent and Soft Computing, Vol. 98)*, Springer-Verlag Berlin/Heidelberg 2012, str. 3-12.

7. A. Burda, P. Cudek, Z.S. Hippe: *ProfileSEEKER – Early Warning System for Predicting Economic Situation of Small and Medium Enterprises*, 6th International Conference on Human System Interaction (HSI'2013), Sopot 2013.