Prof. Ing. Radim Farana, CSc. FEng.
Department of Informatics
Faculty of Business and Economics
Mendel University in Brno
Zemědělská 1
613 00 BRNO
E-mail: radim.farana@mendelu.cz

REVIEW OF THE DISSERTATION THESIS

Title:	Forecasting Regional Financial Performance Using Soft-Computing Methods
Author:	Ing. Evelyn Toseafa
Program:	System Engineering and Informatics
Department:	University of Pardubice Faculty of Economics and Administration Institute of System Engineering and Informatics
Supervisor:	prof. Ing. Petr Hájek, Ph.D.

This review has been written based on the appointment of the opponent by the head of the dissertation committee.

The thesis includes 128 numbered pages including appendices A – J, the bibliography includes 183 references and 8 author's publications. At the time of processing the review, the applicant had 3 results in WoS and 4 in Scopus databases. Number of citations was 1 in WoS and 4 in Scopus databases. However, only three of the publications are cited in the list of references.

The dissertation is a contribution to the forecasting of regional financial performance with the use of soft computing methods and their combinations.

Theme and objectives of the dissertation

The main thesis objective is to apply a novel hybrid model based on the effective combination of different soft computing methods for forecasting regional financial performance in terms of sub-sovereign credit ratings.

This objective is actual and its achievement is not trivial. The main goal was divided into six sub-goals that logically follow each other and end by demonstrating the performance of the developed models.

The thesis presents the achievement of all defined goals. The topic addressed in the thesis is actual and has benefits both for the development of the scientific field and for practice.

Applied methods and results

The dissertation thesis includes a few logically structured chapters presenting separate parts of the work. The first chapter presents an analysis of the state-of-the-art with a great number of citations. The problem domain was analysed from many perspectives. The results of the analysis are concise and to the point. In her analysis, the applicant significantly relies on the results of her supervisor's research and thus expands his scientific school.

The aim and objectives of the dissertation are defined in the second chapter. The main goal was divided into six sub-goals that logically follow each other and end by demonstrating the performance of the developed models.

The third chapter presents the research methodology. It starts with data collection and pre-processing, continued by the selection of significant attributes, training and testing the datasets for machine learning, its realizations, and testing. When presenting an algorithm, I would prefer a classic graphic tool like an Activity Diagram or a Flowchart (compare Fig. 1 and Fig. 3 for example). When using a Flowchart, use it correctly, please (Fig. 3). When presenting algorithms, it is not clear, if they are cited or developed by the author, when the cited source is included in the preceding text only. From Chapter 3.1 I do not understand which datasets were used for the research work, was it years from 2003 to 2016? Or only some of these years?

The fourth chapter describes the experimental settings, applied software tools, their parameters, and application limitations. Could you specify the literature presenting equation (9), please?

Chapter Five presents the obtained experimental results. Four different link functions were tested and seven single classifiers in an experiment for the benchmark dataset with rating classes from 2016 by Moody's credit rating. Five ensemble classifiers were also tested. The performance of feature selection using wrapper and filter feature selection follows. The effect of class balancing using oversampling, ordinal classification, and cons-sensitive classification follows. Random forest model has been evaluated as the best and global SHAP values were used for the model explanation. The chapter concluded with the impact of the variables for all testing data, also showing the direction of the effect. This chapter therefore represents the essence of the presented research.

Limitations and some ideas for further research are presented in Chapter Six. Chapter Seven summarises the contribution of the thesis. Forecasting regional performance is a very complicated requirement, influenced by several state, geographic, and similar restrictions. The use of sophisticated models for decision support is thus always beneficial.

Formal presentation

The presented text is well-structured, included tables and images are well-readable. The structure of the thesis is appropriate.

When presenting an algorithm, I would prefer a classic graphic tool like an Activity Diagram or a Flowchart (compare Fig. 1 and Fig. 3 for example). When using a Flowchart, use it correctly, please (Fig. 3). When presenting algorithms, it is not clear, if they are cited or developed by the author, when the cited source is included in the preceding text only. Similar problems are with some figures, for example, Figure 5, is the author’s work? Numerical scales and units are missing here.

Publications

At the time of processing the review, the applicant had 3 results in WoS and 4 in Scopus databases. Number of citations was 1 in WoS and 4 in Scopus databases. However, only three of the publications are cited in the list of references.

The number of author's publications corresponds to the standard of doctoral studies and demonstrates the student's personal contribution to the problem being addressed. Three author’s publications were listed in the list of references, but only two are cited in the text. Why publication (Toseafa, E. 2017) is referred to but not cited in the text?

All relevant publications are from the years 2017 – 2019. It could help if the student explained why she did not continue in this research and changed the research orientation when the next publications are from the years 2022 – 2023 and are not connected to the thesis.

Questions

1. Artificial intelligence systems are built on existing situations and known values. How do you expect your system to function when evaluating completely new business plans, e.g. built precisely on the massive use of artificial intelligence?
2. The created hybrid model tested only on the data of one company for years from 2003 to 2016, as presented in Chapter 3.1. Do you expect any changes in the evaluation of companies in the post-COVID period, especially caused by the fading energy crisis? How will this be reflected in your models?
3. Is it possible to confront the achieved results with the opinions of experts from practice and with what results?

Final evaluation

I consider the thesis topic to be sufficiently challenging for a dissertation thesis. The author used the correct methods and verified the achieved results. The dissertation thesis can be successfully defended, therefore:

~~ I recommend thesis for the defence ~~

Brno, 2024-03-11

prof. Ing. Radim Farana, CSc. FEng.