

Review of the doctoral thesis:

Data mining in social network analysis

authored by Mgr. Peter Zvirinský

The PhD thesis analyses in detail how insolvency processes work by using social network analysis (SNA) and machine learning, focusing on data from Czech Republic's Insolvency Register. By constructing a dynamic social network of debtors, creditors, and administrators, the study uncovers patterns of debt formation and financial distress, aiming to improve the prediction of insolvency events. This approach not only provides novel insights into the dynamics of financial distress but also offers practical tools for stakeholders involved in managing financial risks. The topic of this study is very important and timely in today's economic world, where financial instability and insolvency have big effects on economies and societies. With financial markets getting more complex and connected, knowing and predicting insolvency processes is key for keeping economies stable.

GOALS:

The goals outlined for the thesis are ambitious and cover a comprehensive approach to analyzing insolvency within the Czech society SNA. Each goal targets a crucial aspect of insolvency data analysis, aiming to deepen the understanding of indebtedness and insolvency dynamics.

METHODOLOGY:

The student demonstrated broad range of technical skills ranging from parsing large volumes of structured and unstructured data efficiently, OCR, data extraction to training the neural network. Student presented wide spectrum of methods for exploratory data analysis, that were able to uncover some patterns in data that were in turn utilized for further investigated. From machine learning point of view, the choice of predictors was very reasonable, even though the motivation for choosing particular algorithm was not always clearly motivated. The student did not consider current state of the art classifiers such as XGBoost, or some modern NN architectures such as TabNet. Overall methodology employed in this work is correct, proves deep knowledge of the topic and ability of the student to perform quality research work.

FORMAL ASPECTS:

The thesis is written in English at a very high level. The student has demonstrated a deep understanding of the topic, with analytical reasoning and strong technical writing skills. The thesis is divided into six main sections. The first two sections provide an overview

of the topic and introduce the reader to the areas of insolvency and social networks. The fourth section offers a detailed description of the dataset and the methods used for data acquisition. The fifth chapter describes the implementation and design considerations. The experiments conducted are thoroughly documented in the sixth chapter, demonstrating the student's ability to apply theoretical knowledge to empirical research. The order and content of the individual chapters are very logical, each building upon the information presented previously.

SUMMARY:

Student employed innovative analysis of social networks, utilizing graph theory, machine learning and data analysis that discovered some interesting observations in insolvency dataset. The thesis provides analytical description of utilized methods and reasonable amount of experiments that were performed and documented on very good scientific level. I believe that these are interesting approaches with relevant results, so it is very unfortunate that have not published research presented in thesis in some reputable international journals. Publication of results is important for community, so others can access or build upon existing methodology and provide additional valuable feedback.

QUESTIONS:

1. Considering the dynamic and temporal aspects of insolvency data, how effective are the chosen machine learning models in capturing temporal dependencies and changes over time? What approaches can be used to improve the incorporation of temporal changes?
2. The research was performed on data from Czech republic. Please discuss
 - a) whether these results are generalizable to similar economics (PL, HU, SK),
 - b) whether the proposed methodology can be applied in different countries.

CONCLUSION:

To conclude, I believe that Peter Zvirinsky has demonstrated a deep knowledge of the topic and the ability to conduct research. As such, I recommend considering his thesis for defense and, if successful, awarding him the degree of Philosophiae Doctor.

prof. Ing. Peter Drotár, PhD.