Report on **Master Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Ing. Petra Kohoutová**
Advisor:	**prof. PhDr. Ladislav Krištoufek, Ph.D.**
Title of the thesis:	**Comparative Analysis of Outlier Detection Models for Transaction Monitoring**

OVERALL ASSESSMENT *(provided in English, Czech, or Slovak):*

Short summary
This thesis provides a comprehensive comparative analysis of three different outlier detection machine learning-based models, particularly focusing on the context of transaction monitoring. Outlier detection is crucial in identifying anomalies in large transactional datasets that might indicate financial fraud, making it an essential tool in fields such as finance and cybersecurity. The research assesses the strengths and weaknesses of these models using both artificial and real-world transactional data. Machine learning plays a key methodological role, particularly in the use of discriminative models, which are preferred to generative models for their performance in large datasets. The thesis specifically explores applicability to one-class classification problems, where identifying rare but impactful events is critical. In conclusion, Isolation Forest and Copulas are the most effective models, balancing the accuracy and computational efficiency trade-off.

Contribution
The thesis provides a very practically oriented topic, elaborated on an industry-standard level. The main scientific contribution lies in the well-designed comparative analysis of a set of machine learning approaches within a specific transaction monitoring context that presents specific challenges, such as missing or unreliable data labels or skewness of the dataset. By focusing on three distinct models, the research critically contrasts and compares their performance, strengths, and weaknesses. This comparison especially extends the understanding of the impact of the dimensionality of the data and the required computational demandingness of the methods, which has not been much explored in previous studies. The results are well-technically described and numerically compared, but I generally miss an attempt to provide the reader with a potential intuitive explanation or a technical interpretation of why some models perform better under specific conditions. Just numerically comparing the metrics is not enough. From a practical point of view, the study offers practitioners clear guidance on which models are most suitable for different types of datasets, particularly when considering high-dimensional and complex environments. The recommendations based on this analysis can directly improve the efficiency and accuracy of current fraud detection systems in finance. The thesis finally suggests more advanced areas for future scientific research, such as evaluating hybrid methods and using larger datasets.

Methods
The thesis evaluates three models: Isolation Forest, K-Means, and a Copula-based approach, examining their sensitivity and specificity in detecting fraudulent transactions. The implementation of the methods is well intuitively motivated and perfectly technically described. The author develops her own artificial testing dataset aside from an empirical dataset consisting of real transactions and designs three increasingly complex scenarios intended to test the models under various conditions. Both the principles of the artificial dataset creation and the technical definitions of the scenarios are described in detail. However, the pseudo-codes of pg. 14 – 16 could have been designed better. Also, the author applies multiple metrics to assess various aspects of the models' performance: the Area Under the ROC Curve (AUC), the anomaly score, and the confusion matrix. The thesis finally evaluates the performance model-by-model for the three scenarios and the two types of data. While model-based ordering makes sense, I would be more interested in a scenario- or data-based ordering and comparison of all models. For instance, all confusion matrices could then be displayed in one figure, providing a more comprehensive overview. This is, nonetheless, partially done in Chapter 7.

Literature
The literature review section supports the motivation of the methodological base of the thesis well. I especially appreciate the well-developed structure, which is focused on important concepts within the field and not on individual papers, a frequent imperfection in many other theses. The number of bibliographic items is not large, which is, however, due to the specificity of the topic. This imperfection is, nonetheless, offset by the extent of the methodological Chapter 3, which introduces another literature associated with the details of individual models. Chapter 6 discusses the presented approach's limitations and practical

Report on **Master Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Ing. Petra Kohoutová**
Advisor:	**prof. PhDr. Ladislav Krištoufek, Ph.D.**
Title of the thesis:	**Comparative Analysis of Outlier Detection Models for Transaction Monitoring**

limitations, appropriately complementing the technical results and helping to communicate the main outcomes.

Manuscript form
The thesis is written in sound English, standardly structured, and typeset in LaTeX. The writing style is very clear and coherent, sometimes a bit technical, but still, the whole work reads well. A (potential) use of generative AI technologies in the writing process should have been declared, but it is clear that the previous year's template used does not contain such a section. The bibliography section is complete and well-formatted. I only spotted one typo for Sklar, A. and another two in the text on pg. 53. The text uses informal contractions "it's" or "Here's" several times. Referencing tables and figures is done correctly in the text, and the tables are well-designed. Some figures, especially in the first part of the text, have a very nice and informative graphics, but later (Figs. 3.4, 3.5, 3.6), the figures would deserve a better description of the content, e.g., in the form of associated notes, making them stand-alone. Figs. 4.1, and 4.3 seem redundant as they only duplicate information already stated in the text. Figs. 4.4 and 4.5 are well-designed, but their content should have been better mentioned in the text; it is not enough to refer to a figure containing so much important information.

Overall evaluation and suggested questions for the discussion during the defense

The assessed thesis clearly surpasses the IES FSV UK master-level standards. Thus, I can confidently recommend it for the defense and suggest a grade of A (excellent).

The results of the Turnitin analysis do not indicate significant text similarity with other available sources.

Additional topics for the discussion:
- The Receiver Operating Characteristic (ROC) Curve and the Area Under the ROC Curve (AUC) are correctly utilized, but their use (although standard in the field) is not sufficiently explained in the text. Can you clarify the intuition and technical details of this performance metric?
- Figure 5.15 presents the K-Means model's qualitatively different (clearly the poorest) performance. Can the author think of and try to explain why that might be so (intuitively or technically based on the specifics of the method)?
- On the other hand, what features make (might make) the winning models the best?
- Chapter 6 mentions practical scenario limitations. What would be the optimal testing scenario in an ideal case of no practical (legal, computational, data-related, etc.) constraints?

SUMMARY OF POINTS AWARDED *(for details, see below):*

CATEGORY		POINTS
Contribution	*(max. 30 points)*	29
Methods	*(max. 30 points)*	29
Literature	*(max. 20 points)*	19
Manuscript Form	*(max. 20 points)*	18
TOTAL POINTS	*(max. 100 points)*	**95**
GRADE	**(A – B – C – D – E – F)**	**A**

NAME OF THE REFEREE: Jiří Kukačka
DATE OF EVALUATION: 10. 9. 2024

_____
Referee Signature

EXPLANATION OF CATEGORIES AND SCALE:

CONTRIBUTION: *The author presents original ideas on the topic demonstrating critical thinking and ability to draw conclusions based on the knowledge of relevant theory and empirics. There is a distinct value added of the thesis.*

METHODS: *The tools used are relevant to the research question being investigated, and adequate to the author's level of studies. The thesis topic is comprehensively analyzed.*

LITERATURE REVIEW: *The thesis demonstrates author's full understanding and command of recent literature. The author quotes relevant literature in a proper way.*

MANUSCRIPT FORM: *The thesis is well structured. The student uses appropriate language and style, including academic format for graphs and tables. The text effectively refers to graphs and tables and disposes with a complete bibliography.*

Overall grading:

TOTAL	GRADE
91 – 100	A
81 - 90	B
71 - 80	C
61 – 70	D
51 – 60	E
0 – 50	F