Report on **Master Thesis**

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

Student:	Bc. Barbora Bajgarová
Advisor:	doc. PhDr. Jozef Baruník, Ph.D.
Title of the thesis:	Machine Learning Methods in Motor Insurance Fraud Detection

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

Short summary

The thesis under review is a rigorous study of how machine learning can be used to detect fraud in motor insurance. It works with a unique real data provided by UNIQUA pojistovna a.s. as well as several algorithms. In addition to comparing modern techniques based on machine learning, the author finds machine learning to be a suitable tool with a high detection rate and reasonable accuracy. The main predictors of fraud, such as the amount of the claim, the type of insurance or the age of the vehicle, are also identified. The results underline the potential of machine learning as a tool to improve the efficiency of fraud detection.

Contribution

The work, which explores the possibilities of using machine learning in detecting fraudulent claims in motor insurance, is novel and contributing in several aspects. It uses a unique and detailed dataset, it compares several advanced techniques with potential for efficient fraud detection, and it provides a comprehensive study of machine learning based techniques in the problem of fraud detection. While this is an important topic for insurance companies, the results are useful as they indicate a significant potential of the machine learning based tools in improving fraud detection. The thesis is also useful from an academic point of view as it contributes to the state of the art through a novel application and empirical results.

Methods

The thesis uses an appropriate set of advanced machine learning tools and compares modern random forests, XBoost, multilayer perceptron combined with SMOTE and class weighting, which are appropriate solutions to the challenge of class imbalance. I must emphasise that Barbora had to deal with a large amount of problems at every step from data preparation itself to estimation, and she was able to deal with them in a professional manner.

Literature

The author uses relevant literature correctly, introduces the literature well and discusses all relevant sources in sufficient detail.

Manuscript form

The thesis is very nicely and clearly written, well motivated and links the findings well with the existing literature. The introduction and motivation are clear and the reader can see that the author has put a lot of care into the text. Barbora has worked consistently on the results over a long period of time and we have discussed the results and the text regularly throughout the year.

Report on **Master Thesis**

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

Student:	**Bc. Barbora Bajgarová**
Advisor:	**doc. PhDr. Jozef Baruník, Ph.D.**
Title of the thesis:	**Machine Learning Methods in Motor Insurance Fraud Detection**

Overall evaluation and suggested questions for the discussion during the defense

In conclusion, I believe that Barbora has shown herself to be an independent researcher who has produced interesting results that are relevant to both the literature and practitioners. She has mastered advanced methods and gained skills to produce a sound and economically relevant work. Barbora has worked consistently with me and I must highlight the hard work she has put into preparing the dataset to be able to answer the questions correctly.

One question to discuss during the defence could be the risks of the proposed solution. In other words, if the results were to be used by an insurance company, what are the practical issues, challenges, and risks that a company would have to deal with?

The thesis deserves to be defended without reservation. If Barbora is confident in presenting the details of the work, I suggest awarding it a grade A.

Finally, the results of the documentary analysis do not indicate any significant similarity of the text to other available sources.

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

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

NAME OF THE REFEREE: Jozef Barunik

DATE OF EVALUATION: 15.8.2024

_____

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