Report on **Master Thesis**

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

Student:	**Bc. Matyáš Mattanelli**
Advisor:	**doc. PhDr. Jozef Baruník, Ph.D**
Title of the thesis:	**Binning numerical variables in credit risk models**

OVERALL ASSESSMENT:

Contribution

The COVID-19 crisis has highlighted the importance of credit risk management and the use of appropriate risk models in finance. However, these models are demanding on the use of data that are not usually available to academics. Fortunately, Matyáš has an access to 5 publicly available data sets, on which he builds his thesis. He has undertaken a thorough analysis of six evaluation metrics, what makes his work outstanding. He also raises several recommendations for the development of credit scoring models. His contribution is therefore clear.

Methods

The author presents 5 research hypotheses. When testing them, he applies standard methods such as employ 5 classification models: Logistic Regression, Decision Tree, Random Forest, Artificial Neural Network and Gaussian Naïve Bayes. Matyáš concludes that the binning transformation has a positive and significant effect on the performance of Logistic Regression, feedforward Artificial Neural Network, and the Naïve Bayes classifier.

Literature

A 9-page literature survey seems comprehensive. Based on this review Matyáš formulates 5 hypotheses. However, some relevant works are not quoted (e.g. Teplý, P., Polena, M. (2020). Best Classification Algorithms in Peer-to-Peer Lending. North American Journal of Economics and Finance) and final results are not compared with other studies.

Manuscript form

The manuscript form fulfils standard requirements. I believe that after some revisions, a part of the thesis might be published in a journal with an impact factor.

Summary and suggested questions for the discussion during the defense

Matyáš has developed a solid academic approach, reviewed the existing literature, identified and investigated open questions and drew relevant conclusions from that. In my view, the thesis fulfills the requirements for a master thesis at the IES, Faculty of Social Sciences, Charles University. I recommend it for the defense and suggest **a grade B.**

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

Questions to the defense:

1) Could Matyáš compare his results with other relevant studies?
2) Why a publicly available data set by the Lending Club, was not included to the thesis? (the author quotes Lessmann et al. (2015), who did research on this data set).

Report on **Master Thesis**

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

Student:	Bc. Matyáš Mattanelli
Advisor:	doc. PhDr. Jozef Baruník, Ph.D
Title of the thesis:	**Binning numerical variables in credit risk models**

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

CATEGORY		POINTS
Contribution	*(max. 30 points)*	27
Methods	*(max. 30 points)*	25
Literature	*(max. 20 points)*	15
Manuscript Form	*(max. 20 points)*	18
TOTAL POINTS	*(max. 100 points)*	**85**
GRADE	**(A – B – C – D – E – F)**	**B**

NAME OF THE REFEREE: prof. PhDr. Petr Teplý, Ph.D.

Digitálně podepsáno (8. 9. 2023)
Petr Teplý

_____

DATE OF EVALUATION: September 8, 2023

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.*

Strong	Average	Weak
30	15	0

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.*

Strong	Average	Weak
30	15	0

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

Strong	Average	Weak
20	10	0

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.*

Strong	Average	Weak
20	10	0

Overall grading:

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