

BUCHAREST UNIVERSITY OF ECONOMIC STUDIES



Doctoral School of Economic Informatics

TITLE OF THE PHD THESIS:

Improvements of business performance using Business Analytics

PhD Supervisor:

Prof. Univ. Dr. BOLOGA Ana Ramona

Thesis Author:

ANTAL-VAIDA Claudia

Bucharest, September 2024

ABSTRACT

The doctoral thesis entitled "Improvements of business performance using Business Analytics" describes the fast-paced and dynamic economic background which has driven major changes in the way industries operate, influencing the adoption of emerging technologies to remain competitive on the market. These trends led to high volumes of data generated each second and to a decreased ability of the manpower to analyze it and use it beneficially. Hence, digital transformation has become a key component of the business strategy, Business Analytics, Artificial Intelligence and Machine Learning becoming more and more popular, but also mandatory for the success of an organization.

The main personal contribution of this paper consists in proposing an analysis methodology based on Machine Learning techniques for data processing in order to obtain information which supports an informed decision-making process and facilitates operational process optimizations, leading to improved performance of the company. This methodology started with the construction of prediction models based on ten Machine Learning classification algorithms which were tested for three public data sets covering problems faced by banks: prediction of default probability of a customer to meet their obligation to the bank, customer churn prediction and fraud detection. In addition to building the models, various optimization methods, such as feature selection, hyperparameters optimizations and resampling techniques, have been included in the proposed methodology and tested to evaluate their performance.

This methodology has a wide range of application areas: it can be used in the banking field, for the areas highlighted in this paper and not only, but also in other fields such as medical, marketing, manufacturing, sports, etc, for classification problems. Its flexible nature can evolve in line with the constantly changing technological trend, ensuring optimal integration of new data processing and analysis techniques.

Keywords: *Business Analytics, Machine Learning, Classification problems, business performance, bank risk management, analysis methodology*

Table of contents

CHAPTER 1. INTRODUCTION.....	1
1.1. Thesis structure.....	3
1.2. Research scope and objectives.....	5
1.3. Reseach methodology.....	6
CHAPTER 2. DIGITAL ERA AND THE EVOLUTION OF BUSINESS ANALYTICS TECHNIQUES	9
2.1. Digital Era.....	9
2.2. Business Analytics.....	12
2.3. Artificial Intelligence.....	13
2.4. Machine Learning.....	16
2.5. Evolution of Artificial Intelligence and Machine Learning.....	18
2.6.Discussions and personal contributions.....	21
CHAPTER 3. CURRENT STATE OF RESEARCH – LITERATURE REVIEW	23
3.1. Trends in credit scoring.....	28
3.2. Trends in customer churn prediction.....	35
3.3. Trends in fraud detection.....	37
3.4. Other trends for addressing banking risks.....	39
3.5. Discussions and personal contributions.....	40
CHAPTER 4. MACHINE LEARNING ALGORITHMS USED FOR RISK MANAGEMENT IN BANKS.....	42
4.1. Risks faced by banks.....	43
4.2. Applicability of Business Analytics and Machine Learning techniques	45
4.3. Difficulties in implementing Machine Learning project	48
4.4. Artificial Intelligence and Machine Learning adoption in banks during the pandemic.....	50
4.4.1. Improving customer experience through conversational banking model	52
4.4.2. Machine Learning usage in Anti-Money Landering System	53
4.4.3. Identity verification to prevent fraud and identity theft	54
4.4.4. Cybersecurity risks	55

4.4.5. SWOT analysis of AI/ML implementation in banks	56
4.5. Discussions and personal contributions	58
CHAPTER 5. BUILDING A CLASSIFICATION MODEL TO- IMPROVE THE STRATEGY OF A COMPANY.....	60
5.1. Algorithms used for classification problems.....	64
5.2. Key performance indicators.....	71
5.3. Improving Machine Learning Algorithms.....	71
5.3.1. Feature selection methods.....	73
5.3.2. Hyperparameters optimization methods.....	77
5.3.3. Resampling methods.....	79
5.4. Discussions and personal contributions.....	80
CHAPTER 6. ANALYSIS METHODOLOGY BASED ON MACHINE LEARNING TECHNIQUES FOR BANKS	81
6.1. Proposed methodology.....	81
6.1.1. Problem definition.....	84
6.1.2. Description of proposed methodology	84
6.1.3. Functions and libraries used to build the models	91
6.1.4. Applicability of the proposed methodology	93
6.2. Dataset description and data processing.....	94
6.3. Model performance indicators.....	103
6.4. Initial performance of the algorithms.....	104
6.4.1. Results analysis for dataset 1.....	104
6.4.2. Results analysis for dataset 2.....	106
6.4.3. Results analysis for dataset 3.....	107
6.4.4. Results interpretation.....	108
6.5. Algorithm performance analysis with feature selection methods.....	108
6.5.1. Feature selection for dataset 1.....	109
6.5.2. Feature selection for dataset 2.....	113
6.5.3. Results interpretation.....	117
6.6. Algorithm performance analysis with hyperparameters optimization.....	118
6.6.1. Results for dataset 1.....	118
6.6.2. Results for dataset 2.....	131

6.6.3. Results interpretation.....	144
6.7. Algorithm performance analysis with resampling techniques.....	144
6.8.Results interpretation.....	154
6.9.Study limitations.....	158
CHAPTER 7. FINAL CONCLUSIONS, PERSONAL CONTRIBUTIONS AND FUTURE RESEARCH DIRECTION.....	159
7.1. Final conclusions.....	159
7.2. Personal contribution.....	167
7.3. Future research directions.....	168
Bibliography.....	169
Annexes.....	192
List of articles and conferences.....	210
List of aronyms.....	214
List of tables.....	216
List of charts.....	219