

# Credit Risk Analysis

Hung Nguyen (1022029) - Hiep Nguyen (1022799)

November, 2022

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation and problem definition . . . . .	1
1.2	Research goals . . . . .	2
<b>2</b>	<b>Data description</b>	<b>2</b>
2.1	Dataset details . . . . .	2
2.2	Data preprocessing and cleaning . . . . .	2
2.3	Explanatory variables selection . . . . .	2
<b>3</b>	<b>Models description</b>	<b>2</b>
3.1	Pooled Logistic Regression model . . . . .	2
3.2	Hierarchical Logistic Regression model . . . . .	2
<b>4</b>	<b>Results analysis</b>	<b>2</b>
4.1	Covergence diagnostics . . . . .	2
4.2	Model comparison . . . . .	2
4.3	Posterior predictive checks . . . . .	2
4.4	Predictive performance assessment . . . . .	2
4.5	Prior sensitivity analysis . . . . .	2
<b>5</b>	<b>Conclusion and potential improvements</b>	<b>2</b>
<b>6</b>	<b>Self-reflection</b>	<b>2</b>
<b>7</b>	<b>Appendices and references</b>	<b>2</b>
7.1	Stan code appendices . . . . .	2
7.2	References . . . . .	2

## 1 Introduction

### 1.1 Motivation and problem definition

Hiep and Hung and Khanh

## 1.2 Research goals

# 2 Data description

## 2.1 Dataset details

## 2.2 Data preprocessing and cleaning

## 2.3 Explanatory variables selection

# 3 Models description

## 3.1 Pooled Logistic Regression model

### 3.1.1 Prior justification

### 3.1.2 Running the model

## 3.2 Hierarchical Logistic Regression model

### 3.2.1 Prior justification

### 3.2.2 Running the model

# 4 Results analysis

## 4.1 Coverage diagnostics

## 4.2 Model comparison

## 4.3 Posterior predictive checks

## 4.4 Predictive performance assessment

## 4.5 Prior sensitivity analysis

# 5 Conclusion and potential improvements

# 6 Self-reflection

# 7 Appendices and references

## 7.1 Stan code appendices

## 7.2 References