Capstone CYO Project Report
Predictiton of Credit Risk for German Bank

Endri Raco

Contents

Dedication

This project and all my work is dedicated to victims of Albanian earthquake 26 November 2019.

Acknowledgement

I would like to express my special thanks of gratitude to Prof. Rafael Irizarry for the wonderful material and thorough explanations he provided during all courses. Also I want to thank my friends of this course who share the same interests for Data Science.

Introduction

Investopedia webpage defines **Credit risk** as the possibility of a loss resulting from a borrower's failure to repay a loan or meet contractual obligations.(Investopedia 2019)

By this definition, we understand that subject who gives the loan (in our case the bank) takes a risk during this process, because it is not known a priori if client who applies and takes the loan will not default (fail to return the money).

Nowadays, banks and other financial institutions are taking advantange of machine learning methods to predict credit risks (refuse loans to risky customers).

The aim of this project is to use machine learning techniques for building a predictive model based on bank client's features. This model should help on understanding which client can be considered as a risky one.

We will use dataset German Credit Risk from Kaggle portal prepared by Professor Dr. Hans Hofmann, University of Hamburg.

In this dataset, each entry represents a person who takes a credit by a bank. Each person is classified as good or bad credit risks according to the set of attributes.

For all project calculations is used the following PC:

```r
print("Operating System:")
```

```
## [1] "Operating System:"
```

```r
version
```

```
##                 _
## platform        x86_64-w64-mingw32
## arch            x86_64
## os              mingw32
## system          x86_64, mingw32
## status
## major           3
## minor           6.1
## year            2019
## month           07
## day             05
## svn rev         76782
## language        R
## version.string  R version 3.6.1 (2019-07-05)
## nickname        Action of the Toes
```

Methods and Analysis

Importing data

Describing Data

Data Wrangling

Exploring Data

Building models

Results

Conclusions

References

Investopedia. 2019. "What Is Credit Risk?" May. https://www.investopedia.com/terms/c/creditrisk.asp.