From Payment Services Directive 2 (PSD2) to Credit Scoring: A Case Study on an Italian Banking Institution

Roberto Saia, Alessandro Giuliani,
Livio Pompianu, Salvatore Carta

Department of Mathematics and Computer Science
University of Cagliari, Italy

KDIR 2021, 13th International Conference on Knowledge Discovery and Information Retrieval
25÷27 October 2021 – Online Streaming

Outline

- **Overview**
  Introduction ▪ State of the Art ▪ Open Problems
- **Idea**
  Intuition ▪ Formalization
- **Implementation**
  Data Aggregation ▪ Meta-features Addition ▪ Classification
- **Results**
  Experiments ▪ Conclusions

OVERVIEW

Introduction

- The main aim of **credit scoring models** is the classification of loan applications (from now on named as ***instances***) as *reliable* or *unreliable*

Introduction

- They are usually **designed** on the basis of the past loan applications
- They represent a **crucial** instrument for financial operators

State of the Art

- Many **techniques** have been designed to address the *Credit Scoring* problem, by exploiting several models, such as:
  - *Statistical* and *Data Mining*
  - *Linear Discriminant*
  - *Logistic Regression*
  - *Neural Networks*
  - *Genetic Program*
  - *Decision-tree*
  - … and many more

Open Problems

- Regardless of the adopted technique, the definition of effective *Credit Scoring* models is a **hard challenge** for several factors

- The most important of which is the **Imbalanced class distribution**, i.e., a number of *non-refunded loans* **much smaller** than that of the *refunded ones*

- This reduces the **effectiveness** of the classification approaches

RELIABLE CASES **VS** UNRELIABLE CASES

IDEA

Intuition

- To mitigate the aforementioned **open problems,** we introduced two elements:

  1. The opportunities offered by the **Payment Services Directive 2** (PSD2), a new regulatory framework introduced by the European Commission

  2. An **extended feature space** defined by adding a series of meta-information

Formalization

- The PSD2 regulation allows third parties to obtain free access to the client accounts and their **payment transactions** through bank APIs
- This allows us to evaluate user's reliability **without using the canonical information** (e.g., age, total incomes, previous loans, etc.)
- In this way, the financial services that need a solvency assessment can be **extended** to a wide range of users
- We also introduce a feature extension aimed to **better characterize** each user' instance

IMPLEMENTATION

Implementation

Step 1 of 3 - Data Aggregation Approach (DAA)

In order to train the classification model, as first step we aggregate the transaction information of each user in a single data vector on the basis of the following seven criteria:

- number of transactions
- activity days
- minimum handled amount
- maximum handled amount
- total handled amount
- mean handled amount
- standard deviation of all the transactions

Implementation

Step 2 of 3 - Meta-features Addition Approach (MAA)

The data features are then extented by adding to them a series of meta-features MF: **Minimum** (mf_1), **Maximum** (mf_2), **Average** (mf_3), and **Standard Deviation** (mf_4)

$$MF = \begin{cases} mf_1 = min(f_1, f_2, \dots, f_N) \\ mf_2 = max(f_1, f_2, \dots, f_N) \\ mf_3 = \frac{1}{N} \sum_{n=1}^{N} (f_n) \\ mf_4 = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} (f_n - \bar{f})^2} \end{cases}$$

Implementation

Step 3 of 3 – Classification

Each instance i is then classified as **reliable** or **unreliable**, in accordance with the criteria formalized in the following **algorithm**:

Algorithm

Require: A=Classifier, I=Set of classified instances, $\hat{i}$=Instance to evaluate
Ensure: c=Classification of the $\hat{i}$ instance

procedure GETEVALUATION(A, I, $\hat{i}$)
- $I \leftarrow aggregateData(I)$ ▷ DAA process on I set
- $\hat{i} \leftarrow aggregateData(\hat{i})$ ▷ DAA process on $\hat{i}$ instance
- $I \leftarrow addMetafeatures(I)$ ▷ MAA process on I set
- $\hat{i} \leftarrow addMetafeatures(\hat{i})$ ▷ MAA process on $\hat{i}$ instance
- $model = trainModel(\mathrm{A}, I)$ ▷ Classification model training
- $c = getPrediction(model, \hat{i})$ ▷ Instance classification
- **return** c

end procedure

RESULTS

Experiments

- The validation has been performed using **two real-world datasets** provided by a huge Italian bank:
  - The first dataset, named *Credit Card Transactions* (**CCT**), contains **transactions** (3,376,573) related to the use of a revolving credit card
  - The second dataset, named *Bank Customers Information* (**BCI**), refers to the same users in the CCT dataset (49,847), and it contains the **canonical** Credit Scoring information usually exploited in the literature

Experiments

- The experiments involve five of the most performing credit scoring algorithms: *Gradient Boosting* (**GB**), *AdaBoost* (**AB**), *Random Forests* (**RF**), *Multilayer Perceptron* (**MP**), and *Decision Tree* (**DT**)

- The performance comparison has been made by taking into account the average value of three metrics: **Sensitivity**, **Specificity**, and **AUC**

Experiments

Canonical Credit Scoring Model Performance:

Algorithm	Dataset	Sensitivity	Specificity	AUC	Average
GB	BCI	0.9320	0.9295	0.9307	0.9307
AB	BCI	0.8933	0.8815	0.8873	0.8874
RF	BCI	0.9560	0.9992	0.9776	0.9776
MP	BCI	0.4836	0.7899	0.6367	0.6367
DT	BCI	0.9613	0.9712	0.9715	0.9715

Transactions-based Credit Scoring Model Performance, Before and After the Meta-features Addition:

Algorithm	Dataset	Average before	Average after	Increment (%)
GB	CCT	0.8190	**0.8208**	+0.22%
AB	CCT	0.7013	**0.7056**	+0.62%
RF	CCT	0.8980	**0.9118**	+1.54%
MP	CCT	0.6233	**0.6354**	+1.94%
DT	CCT	0.9371	**0.9397**	+0.28%

Conclusions

- the adoption of the **canonical credit scoring model** leads toward better performances than those obtained using only the bank transactions

- But the **transactions-based credit scoring model** offers the opportunity to extend the set of potential users

- In addition, the **introduction of simple meta-features** improves the credit scoring performance of all the classification algorithms

- The meta-features improvement could appear slight, but we can consider it **an interesting result** in the real-world scenarios, where usually are involved a huge number of users

From Payment Services Directive 2 (PSD2) to Credit Scoring: A Case Study on an Italian Banking Institution

THANK YOU FOR YOUR ATTENTION