

Prédiction de la souscription à un dépôt à termes

Karim BENKIRANE

March 16, 2018

Abstract

Les données utilisées sont tirées d'UCI machine learning repository. Il s'agit des données "Bank Marketing Data Set" publiées en 2012. Nous avons été confronté à un déséquilibre de notre échantillon que nous avons dû corriger; aussi, une stratégie de discrétisation des variables a été mise en place avant de lancer un Random Forest et une SVM.

1 Introduction

Le jeu de données dont nous disposons résulte d'une action de marketing direct menée par une banque portugaise. Les prises de contacts avec les prospects se font par téléphone et requiert parfois plus d'un appel avec la même personne. Le but de cette campagne est de promouvoir la souscription à un dépôt à termes. Elle a permis de collecter un certains nombres de données sur les prospects avec pour résultat la souscription ou non au produit bancaire. Nous tâchons donc dans le rapport qui suit de prédire la souscription à ce dernier. Dans un premier temps nous procédons à une brève description des données (caractérisée notamment par un dsiquilibre), avant de passer au retraitement des données (discrétisation notamment) pour finir par la construction des modèles Random Forest, SVM) et le résultat des prédictions.

2 Preprocessing

2.1 Description générale de la table

La table contient 41188 observations représentant un prospect et 21 variables dont 11 catégorielles et 10 quantitatives. La variable à prédire étant y (souscription à un dépôt à termes ou non) Pour en citer certaines - Catégorielles: le type d'emploi, mois, le jour... - Discrètes: âge, nombre de fois contacté au cours de la campagne... - Continues: indice de confiance du consommateur...

Par ailleurs la table ne contient aucune observation avec au moins une donnée manquante. Intéressons nous à la répartition du label y (accepter de souscrire à un dépôt à termes ou non)

2.2 How to add Comments

Comments can be added to your project by clicking on the comment icon in the toolbar above. To reply to a comment, simply click the reply button in the lower right corner of the comment, and you can close them when you're done.

Comments can also be added to the margins of the compiled PDF using the todo command, as shown in the example on the right. You can also add inline comments:

This is an inline comment.

Here's a
comment
in the
margin!

2.3 How to add Tables

Use the table and tabular commands for basic tables — see Table 1, for example.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.0	102.0	180.0	258.3	319.0	4918.0

Figure 1: Statistiques descriptives de la durée du contact téléphonique.

Item	Quantity
Widgets	42
Gadgets	13

Table 1: An example table.

2.4 How to write Mathematics

L^AT_EX is great at typesetting mathematics. Let X_1, X_2, \dots, X_n be a sequence of independent and identically distributed random variables with $E[X_i] = \mu$ and $\text{Var}[X_i] = \sigma^2 < \infty$, and let

$$S_n = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{1}{n} \sum_i^n X_i$$

denote their mean. Then as n approaches infinity, the random variables $\sqrt{n}(S_n - \mu)$ converge in distribution to a normal $\mathcal{N}(0, \sigma^2)$.

2.5 How to create Sections and Subsections

Use section and subsections to organize your document. Simply use the section and subsection buttons in the toolbar to create them, and we'll handle all the formatting and numbering automatically.

2.6 How to add Lists

You can make lists with automatic numbering ...

1. Like this,
2. and like this.

... or bullet points ...

- Like this,
- and like this.

2.7 How to add Citations and a References List

You can upload a `.bib` file containing your BibTeX entries, created with JabRef; or import your [Mendeley](#), CiteULike or Zotero library as a `.bib` file. You can then cite entries from it, like this: [\[Gre93\]](#). Just remember to specify a bibliography style, as well as the filename of the `.bib`.

You can find a [video tutorial here](#) to learn more about BibTeX.

We hope you find Overleaf useful, and please let us know if you have any feedback using the help menu above — or use the contact form at <https://www.overleaf.com/contact>!

References

- [Gre93] George D. Greenwade. The Comprehensive Tex Archive Network (CTAN). *TUGBoat*, 14(3):342–351, 1993.