



ÅBO AKADEMI UNIVERSITY

MACHINE LEARNING

Banking Campaign Output Prediction



FILIBE FELÍCIO (2004624)

JANUARY 30, 2021

Contents

0.1	Introduction	3
0.2	Data Processing	3
0.2.1	Data-set Loading	3
0.2.2	Processing	3
0.3	Modelling	4
0.4	Conclusion	4

0.1 Introduction

This task is to train a Neural Network Model to predict whether a client would respond positive or negative to the campaign.

The features contained in the dataset are as follow:

- age (numeric)
- job : type of job (categorical)
- marital : marital status (categorical)
- education (categorical)
- default: has credit in default? (categorical)
- housing: has housing loan? (categorical)
- loan: has personal loan? (categorical)
- contact: contact communication type (categorical)
- month: last contact month of year (categorical)
- day_of_week: last contact day of the week (categorical)
- duration: last contact duration, in seconds (numeric)
- campaign: number of contacts performed during this campaign and for this client (numeric)
- pdays: number of days that passed by after the client was last contacted from a previous campaign (numeric; 999 means client was not previously contacted)
- poutcome: outcome of the previous marketing campaign (categorical)
- emp.var.rate: employment variation rate - quarterly indicator (numeric)
- cons.price.idx: consumer price index - monthly indicator (numeric)
- cons.conf.idx: consumer confidence index - monthly indicator (numeric)
- uribor3m: euribor 3 month rate - daily indicator (numeric)
- nr.employed: number of employees - quarterly indicator (numeric)
- y (desired target): has the client subscribed a term deposit? (binary)

0.2 Data Processing

0.2.1 Data-set Loading

I used the method `read_csv` from the library `pandas` to load the `bank-additional-full.csv` that was given.

0.2.2 Processing

For the Data Processing I decided to apply One Hot Encoding to the non-numerical features using the function `OneHotEncoder` from the `sklearn.preprocessing` library. After I chose to split the Data-set into two Sets with 66.6% and the remaining 33.3% for Training and Testing, respectively, applying the method `train_test_split` from the `sklearn.model.selection`.

0.3 Modelling

For the Model I decided for a Neural Network with a Sequential model from the keras library.

0.4 Conclusion

For me, the main challenge was to figure first which was the best approach from all the machine learning algorithms that were given in the theoretical part of the class. Then after learning the high-level libraries that abstracted the algorithm I desired for me to use to solve the particular task that was given.