Intro to ML: Coursework Part II

M.Sc. in Financial Englineering
Master Level M2 - Academic Year 2024-2025
Professor: Prof. D. O'Kane

Deadline:

9.00AM Monday 4th November 2024

1   Introduction

This coursework consists of a single case study in which you seek to build the best model you can to predict the default of a loan. These are peer-to-peer loans covering the period from 2007-2017 and have been provided by Lending Club. You need to predict the label 'loan_status' having a value called 'charge-off' - a charge-off refers to debt that a company believes it will no longer collect as the borrower has become delinquent on payments.

The dataset consists of 100,000 loans and 151 features (including the 'charge-off' column). Of these loans, you will find that only 19% have been charged off. Your job is to use the skills you have learned to build the best possible model to predict the defaulted loans.

Note that you will need to use the GridSearch and Feature Importance features of Scikit learn in order to choose the best hyper parameters. Read the Scikit Learn documentation at `https://scikit-learn.org/stable/` to see how to use them.

To guide you on the work, I have set out detailed instructions in the provided notebook. However you should see if you can do something original that will increase your score. Finally, imagine that you are writing the notebook to explain very clearly what you have done to your boss in your next job ! Good luck !

2 Coursework Instructions

Please read these and ensure that you conform.

- You must be in groups of 3 or 4.

- This coursework forms 15% of your overall course grade.

- Download and unzip the coursework file from the Blackboard

- There is this document, a notebook, a data dictionary and a zipped data file that should be unzipped.

- Use sub-headings to divide up the sections and write clear explanations explaining your decisions and results (do not be too verbose). Make sure that all graphs are clearly labeled. You will lose points if your presentation is poor.

- You must work as a team.

- The penalty for copying is to get zero.

3 IMPORTANT - Submission Instructions

Please follow these instructions carefully.

- One person in each time must submit a zip file (.zip only) containing the notebook.

- Email the zip file to me at dominic.okane@edhec.edu before the deadline.

- Ensure that the 3 or 4 student IDs are in the name of the zip file. The filename should be ML_Coursework_Two_Submission_WWWWW_XXXX_YYYY_ZZZZ.zip where WWWWW, XXXXX, YYYYY and ZZZZZ are the student IDs of the students in your team.