Final Projects for FIE 453 fall 2019

October 4, 2019

Each group has to submit the final project as part of FIE 453. The submitted project will be assessed based on both execution and the presentation. Empirical assignments should be solved in R. Results shall be presented in a report no more than 20 pages as well as a PowerPoint presentation that will be presented as part of the project. The corresponding R file(s), PowerPoint presentation and report in pdf shall be submitted to Wiseflow before the deadline 11. November 2019. Note that your work should be reproducible, meaning all code should be runnable. Using R Markdown will ensure this. Otherwise you will have to add a description on how to run all the code needed for reproducing your analysis and results.

The dataset used in this assignment consists of features and a response variable for companies at a given date. The response variable is in the first column and describes whether each company went into default or not in the following 12 months. The remaining columns consists of financial information, company information and behavior data related to each company.

Assignment 1
Develop a credit risk model that predicts the probability of default for the companies in this dataset (behavior model). You should develop three different algorithms– a logistic regression model, a xgboost model and a random forest model. When developing the model ensure that you apply prudent methods of model development in order to handle the bias-variance tradeoff. Describe:

- The models you have developed as well as strengths and weaknesses with the chosen methods.
- How you have ensured a prudent model development framework (train-test method)
- Differences in prediction between two of the models you have developed and try to identify the largest deviations in the models (in terms of predictions).

Assignment 2
Evaluate the model performance of the three models and rank them. Describe:

- Which performance measures you apply and why you apply them.
- Which other “business” assessments should be carried out in addition to statistical validation tests to assess model performance
- How you would assess model performance if the portfolio had been a low-default portfolio

Assignment 3
Assume you are developing a credit risk model for retail customers. How can data stemming from PSD II be utilized in a PD model for retail customers?

Assignment 4
Give a short description of GDPR and how this regulation might impact model development within the financial sector?

Assignment 5

Calculate the risk weighted asset for each customer according to the CRD (see https://www.bis.org/publ/bcbs128b.pdf - page 64).

What is the interpretation of RWA?

Which models yields the lowest RWA and why is that?

Hint! The latter relates to the construction of the RWA formula

$$\text{Correlation } (R) = 0.12 \times (1 - \text{EXP}(-50 \times PD)) / (1 - \text{EXP}(-50)) + 0.24 \times [1 - (1 - \text{EXP}(-50 \times PD)) / (1 - \text{EXP}(-50))]$$

$$\text{Maturity adjustment } (b) = (0.11852 - 0.05478 \times \ln(PD))^2$$

$$\text{Capital requirement}^{72} \ (K) = [LGD \times N[(1 - R)^{-0.5} \times G(PD) + (R / (1 - R))^{0.5} \times G(0.999)] - PD \times LGD] \times (1 - 1.5 \times b)^{-1} \times (1 + (M - 2.5) \times b)$$

$$\text{Risk-weighted assets (RWA)} = K \times 12.5 \times EAD$$

We assume in this assignment for all exposures that LGD (loss given default) equals 45 %, EAD = 100, M (maturity) = 2,5 years while S (size) = 50.

Assignment 6

Assume that you have access to all the companies accounting systems (such as Visma, Uni Micro, PowerOffice etc.) through an API. What information could potentially be used in the PD models that you developed in assignment 1?

Description of variables in the dataset:

Description	Variable name
Default	Mislighold
Profit margin	Resultatgrad
Gross operating income (percent)	Bruttofortjeneste_prosent
Operating margin	Driftsmargin
EBITDA margin	EBITDA_margin
Interest coverage ratio	Rentedekningsgrad
Cost of debt	FremmdekapitalKostnad
Interest bearing debt / EBITDA	Rentebærende_gjeldEBITDA
Revenue stability	Stabilitet_omsetning
Equity ratio	EK_andel
Equity ratio stability	EK_andel_stabilitet
Current ratio	Likviditetsgrad1
Quick ratio	Likviditetsgrad2
Liquidity ratio 3 (cash/current liabilities)	Likviditetsgrad3
Equity	Størrelse_EK
Total assets	Størrelse_balanse
Revenue	Størrelse_omsetning
Age of company	Alder
Unpaid debt collection (sent to debt collection)	Uoppgjorte_inkasso
Paid debt collection (sent to debt collection)	Oppgjorte_inkasso
Adverse opinion (audit)	Revisoranmerkning
industry	Bransje
Amount unpaid debt collection	Belop_uoppgjort_inkasso
Payment reminders	Purringer