

MAFS6010Z: Project1 Review (Group 9)

Summary: Group 9 applies Linear Discriminant Analysis to all independent variables and dependent variable namely TARGET in the training set to analyze the default probability of each client. The result shows that such model has an accuracy of 0.75 in the training set and an accuracy of 0.68 in the test set.

Strengths of the report: The model, training results and conclusion are written in a concise and clear form.

Weaknesses of the report: The reliability of part of the data preprocessing remains questionable and the selection of decision probability is lack of clarification.

Evaluation on Clarity and quality of writing: I would give the report 5 on 5 as it's written flawlessly.

Evaluation on Technical Quality: I would give the report 4 on 5. In part of the codes that belong to data preprocessing, group 9 treats the unique categories of each categorical variables as series of unique ordinals which may increase the inaccuracy of model prediction since we don't know which category of a given categorical variable will bring bigger impact on default probability in advance. In addition, the report doesn't give an explanation of choosing 0.1 as the optimal decision probability.

Overall rating: I would give the report 4 on 5.

Confidence on my assessment: 3 - I have carefully read the paper and checked the results.