



Hackathon Business Case

Business Case & Challenges

25Credit (name has been changed to protect identity) is a global credit card company. Recently, they had experienced \$10 million of fraudulent transactions annually. Many of these fraudulent transactions were not detected by their traditional rules based identification engine and were only reported when customers found unauthorized transactions reflected on their monthly bills.

This has caused the company significant costs and effort in the investigation and management of these transactions due to the lag time from the occurrence of the fraud to the reporting. In addition, it was hard to recover the full amount from those transactions due to the time lapse.

As a result, 25Credit is looking for solutions that will help them reduce the time and effort for investigations as well as improve on the fraud recover through early identification and proactive management of suspicious transactions.

Task for your team

Your team has been brought in as consultants to solve the business challenges of 25Credit. We will be expecting you to share your analysis and investigation into the current problem and work out a business solution proposal including the following:

- Analysis & understanding of the current problem & data
- Technical ML model based on the real life dataset provided
- Recommendations on integrating the insights into action plans / solutions to tackle 25Credit's business challenges

About the data set

A dataset containing actual real life transactions of 25Credit will be shared for analysis and the building of machine learning algorithms. Due to confidentiality, sensitive features have been protected through PCA transformation into its principal components.

The dataset contains:

- Column 1: No. of elapse seconds since reference time
- Column 2: Transaction Amount in USD
- Column 3: Classification (1 for fraud, 0 for non-fraud)
- The remaining columns: PCA feature 1 to 29
- Addition information: Dataset contains 227844 transactions, of which 417 are classified as fraud

Expected Submission materials

1. Business Proposal Report in .pptx format. The following needs to be covered:
 - Clear visual analysis, graphs & charts to show understanding on dataset / features
 - Model & Methodology of Machine Learning Model used
 - Accuracy measure of ML model and justification of metric selection
 - Solution proposals on how business should employ the ML model and insights in an operational context
 - Any assumptions used for formulation
 - Please keep the report to a maximum of 30 slides. Any slides above that will not be assessed.
2. Python Notebook in .ipynb format
 - All graphs, charts, model and accuracy score needs to be shown in the submitted Jupyter notebook
 - Comments and Insights in the notebook will **NOT** be taken into account and these should be incorporated in the PowerPoint report

Judging Criteria

Overall Analysis & Presentation

30%

- Relevant & appropriate use of visualizations to enhance story telling
- Crisp & concise presentation of findings and insights
- Logically structured arguments and conclusion
- Clear explanation of concepts in simple terms

ML Methodology & Model

50%

- Sound data transformation & data preparation methods employed
- Appropriate selection and application of model to dataset
- Decision of parameter values employed
- Proper accuracy metrics used with justifications
- Appropriateness & reasonable assumptions
- Solution & model needs to be reproducible and will be evaluated against a test set
- Clean coding

Business Solution Proposal

20%

- Feasible & realistic recommendations for application of model
- Innovation & creativity in application of insight to real life business context

Presentation Day

Selected semi finalists & finalists will be invited to Cognizant office to meet our judges who are veterans in the data analytics industry. During the presentation, students will be required to prepare:

1. A concise set of presentation materials to highlight on the main points
2. Live demonstration of their codes & models on the IBM private cloud platform

Additional areas of assessment during presentation

In addition to the judging criteria, selected teams who made it to the presentation rounds will also be assessed on the following during presentation:

- Overall presentation of business proposal
- Logical flow in analysis & conclusion
- Effectiveness of communication
- Visual aids and use of visualization
- Completeness of responses to judges' questions