

JJMALK Data Insights, Inc.

Model for Predicting Credit Card Fraud for BOMI

Milestone Status Report

Project Name: Model for Predicting Credit Card Fraud

Reporting Period: December 13, 2023

Project Manager: Maricris Resma

Project Owner: JJMALK Data Insights, Inc.

Executive Summary

Overall Status: **ON TRACK**

| | Status | Reason for Deviation |
|----------|----------|--|
| Budget | ON TRACK | n/a |
| Schedule | ON TRACK | Buffer was used but still ON TRACK for final milestone |
| Scope | ON TRACK | n/a |

Comments:

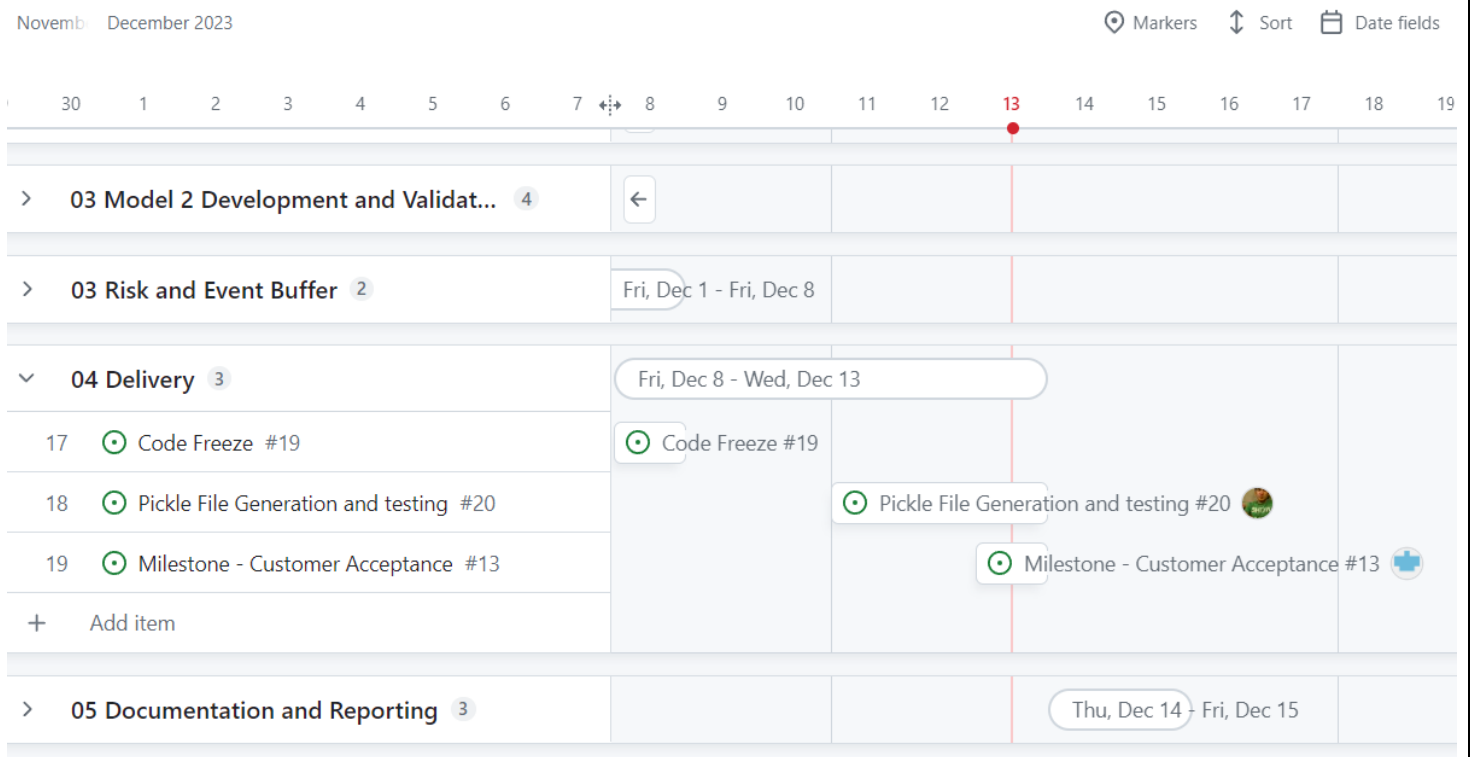
- The first Project Phase: Planning has been done
- The project **Milestone - Customer Acceptance** has been completed on December 13, 2023

Project Definition

| | |
|---------------------------|---|
| Project Objectives | <ul style="list-style-type: none"> Develop a predictive modelling application that effectively predicts credit card fraudulent transactions |
| Scope | <ul style="list-style-type: none"> Retrieve and manage credit card transaction Data from Google Cloud Platform Create a machine learning classifier model capable of detecting whether credit card transaction is fraud or not(minimum 80% accuracy, recall and f1 score) using Decision Tree as the initial Model then Random Forest in case initial model does not reach desired accuracy. Use sklearn and imblearn libraries. Choose Model with best balance between reduced false positives to guarantee a seamless and trustworthy user experience but more importantly high precision to protect from financial loss Deliver the pickle file that can be integrated into Bank of Mississauga's organizational systems and databases |
| Assumptions | <ul style="list-style-type: none"> Customer BOMI (Bank of Mississauga) will be responsible for integrating the delivered pickle file in their system |

Scheduled Milestones / Deliverables

Gantt Chart of Project Timeline



Project Timeline and Deliverables

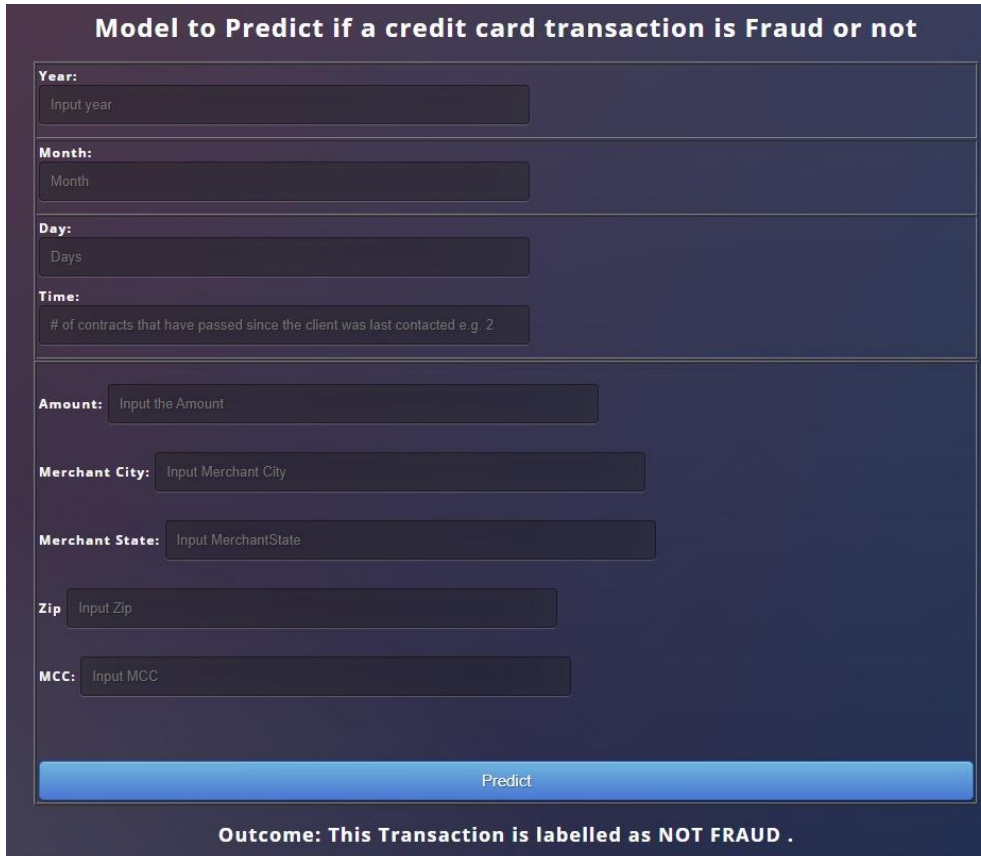
| Project Tasks | Start | End | Deliverables | Status | Actual |
|---|-------------|--------------------|---|-------------|--------------------|
| Planning | | | | | |
| Proposal Creation and Submission | 06-Oct-2023 | 12-Oct-2023 | | DONE | 12-Oct-2023 |
| Data Gathering | 13-Oct-2023 | 20-Oct-2023 | | DONE | 20-Oct-2023 |
| Model Research and Design | 13-Oct-2023 | 20-Oct-2023 | | DONE | 20-Oct-2023 |
| ◆ Milestone- Planning Done | | 20-Oct-2023 | Proposal Doc | DONE | 20-Oct-2023 |
| Model 1 Development and Validation | | | | | |
| GIT Creation | 23-Oct-2023 | 23-Oct-2023 | | DONE | 23-Oct-2023 |
| Data Preprocessing | 24-Oct-2023 | 03-Nov-2023 | | DONE | 03-Nov-2023 |
| Exploratory Data Analysis | 06-Nov-2023 | 10-Nov-2023 | | DONE | 10-Nov-2023 |
| Data Modeling Using ML Algorithm 1 | 13-Nov-2023 | 22-Nov-2023 | | DONE | 22-Nov-2023 |
| Model 1 Performance Evaluation | 13-Nov-2023 | 22-Nov-2023 | | DONE | 22-Nov-2023 |
| ◆ Milestone - Model 1 Dev and Val Done | | 23-Nov-2023 | | DONE | 23-Nov-2023 |
| Model 2 Development and Validation (if needed) | | | | | |
| Data Modeling Using ML Algorithm 2 | 24-Nov-2023 | 28-Nov-2023 | | DONE | 28-Nov-2023 |
| Model 2 Performance Evaluation | 24-Nov-2023 | 28-Nov-2023 | | DONE | 30-Nov-2023 |
| Model Selection | 29-Nov-2023 | 30-Nov-2023 | | DONE | 8-Dec-2023 |
| ◆ Milestone - Model 2 Dev and Val Done | | 30-Nov-2023 | | DONE | 8-Dec-2023 |
| Delivery | | | | | |
| Code Freeze | 01-Dec-2023 | 01-Dec-2023 | | DONE | 12-Dec-2023 |
| Pickle File Generation and Delivery | 04-Dec-2023 | 04-Dec-2023 | | DONE | 13-Dec-2023 |
| ◆ Milestone - Customer Acceptance | | 04-Dec-2023 | zip file of .ipynb and pickle file | DONE | 13-Dec-2023 |

| Project Tasks | Start | End | Deliverables | Status | Actual |
|------------------------------------|-------------|--------------------|----------------------------------|-------------|------------|
| Documentation and Reporting | | | | | |
| Documentation | 05-Dec-2023 | 07-Dec-2023 | | ONGOING | |
| Create Presentation | 05-Dec-2023 | 07-Dec-2023 | | NOT STARTED | |
| ◆ <i>Milestone - Presentation</i> | | 08-Dec-2023 | <i>.doc report and .ppt file</i> | NOT STARTED | |
| Risk and Event Buffer | | | | | |
| Risk Buffer | 11-Dec-2023 | 13-Dec-2023 | | DONE | 1-Dec-2023 |
| Event Buffer | 14-Dec-2023 | 15-Dec-2023 | | DONE | 7-Dec-2023 |

Note: ◆ signify Key Milestones

Accomplishments during this Reporting Period:

1. Code Freeze
 - Code Label ini Git done
2. Pickle File Generation
 - 8 pickle files generated for data transformation and decision tree model prediction
3. Create basic UI to test new data



Model to Predict if a credit card transaction is Fraud or not

Year:
Input year

Month:
Month

Day:
Days

Time:
of contracts that have passed since the client was last contacted e.g. 2

Amount: Input the Amount

Merchant City: Input Merchant City

Merchant State: Input MerchantState

Zip Input Zip

MCC: Input MCC

Predict

Outcome: This Transaction is labelled as NOT FRAUD .

4. Model Deployment

Next Steps

Plans for the next Reporting Period:

Milestone - Presentation

1. Create Validation Test Cases Matrix
2. Create PPT Presentation
3. Demo

Issue Status:

- No Issues occurred

Change Status:

- No change needed

Risk Status:

| Risk | Impact Level | Mitigation Plan | Status |
|--|--------------|--|--------|
| Imbalanced data results in a data model that may lead to bias toward the majority class and poor predictions for the minority class | High | Apply resampling to make up for the imbalanced data | Closed |
| Poor accuracy in data model due to poor quality of data may lead to difficulty that can cause delay | Medium | Apply rigorous data preprocessing to ensure removal of outliers and irrelevant data | Closed |
| The resulting performance metrics of initial modelling algorithms to be used may not reach required accuracy, causing delay in remodelling | Low | Extensive pre-study on the data and research on best model to use must be done beforehand | Closed |
| There is a chance that the project won't be completed in the allotted time due to a team member's illness | Low | The team will apply effective time management by developing efficient work breakdown where there are primary and secondary members assigned to each task | Closed |

Notes: Status Definition

Planned - An identified risk with a risk response plan.

In-Process - A risk where the risk response is being executed

Closed - A risk that occurred and is transferred to an issue or the risk was solved/avoided

Not occurred - A risk that was identified but that did not occur

Rejected - Created and kept for tracking purposes but considered not to be used yet

Budget Status as of 20/10/2023

| Budget Category | Original Budget Cost | Spent to Date | Budget Left |
|---------------------------------------|----------------------|-----------------|----------------|
| Software Costs | | | |
| Application software | \$500 | \$500 | \$0 |
| Operating system | \$1,000 | \$1,000 | \$0 |
| Hardware Costs | | | |
| PCs | \$10,000 | \$10,000 | \$0 |
| Network Costs | | | |
| Internet | \$1,000 | \$1,000 | \$0 |
| Team Resources | | | |
| Project Manager (\$30/hr) (\$240/day) | \$12,240 | \$9,000 | \$2,240 |
| Technical Lead (\$30/hr) | \$12,240 | \$9,000 | \$2,240 |
| Developer (\$30/hr) | \$12,240 | \$9,000 | \$2,240 |
| Researcher (\$30/hr) | \$12,240 | \$9,000 | \$2,240 |
| Validator (\$30/hr) | \$12,240 | \$9,000 | \$2,240 |
| Total | \$ 73,700 | \$70,550 | \$3,150 |

Budgetary Comments:

- Still within budget