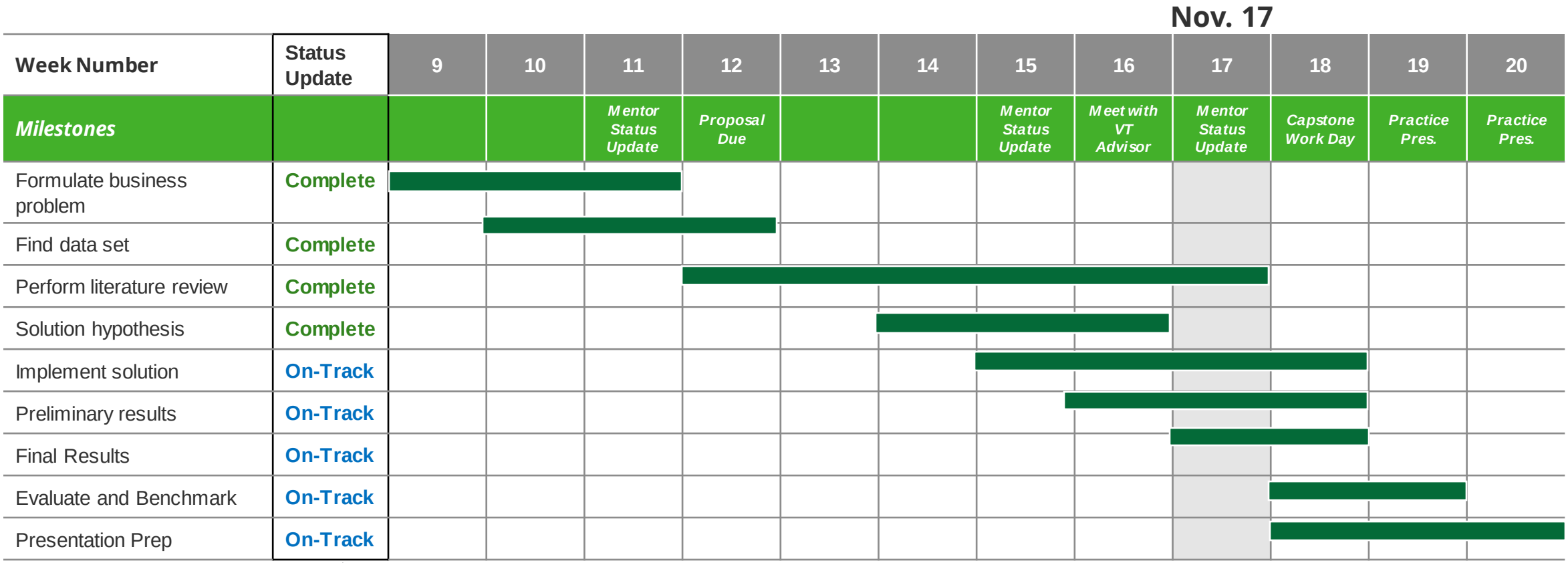
Project Timeline

November 17, 2023



Indicate:
Complete
On Track
At Risk
Off Track

Capstone Proposal: Bank Account Fraud Detection

Project Name	Bank Account Fraud Detection
Business Understanding	Our team is aiming to detect bank account fraud detections by evaluating a variety of variables related to financial status, personal identification, and payment patterns. We will uncover the answers to questions such as: 1. Which factors should be prioritized when monitoring bank account fraud? 2. Which predictors affect fraud detection the most? 3. Are there trends for the frequency of fraud types?
Client/Project Motivation	Our clients are banks and financial institutions. Bank account fraud costs companies millions of dollars every year and creating a proactive approach towards bank account fraud detection may help cut costs and increase awareness of fraud factors or triggers.
Market /Industry	Finance, Banking, Insurance
State of the Art	Our team plans to make our model state of the art by incorporating logistic regression and building out additional features that other models are currently not considering.
Success Metrics (Evaluation)	Our team will evaluation success by examining precision, recall, confusion matrix, accuracy rates, ROC and AUC scores.
Scalability	The team will build scalability into the model throughout the development stage by experimenting with data set size and varying feature counts to increase the likelihood of continued effectiveness. Employing Logistic Regression as well will improve performance.

Modeling Requirements

Data Type	CSV file: numeric & categorical variables
Data Source	**Kaggle:** Bank Account Fraud
Data Preparation Steps	Look at distributions, patterns and correlations. Apply appropriate filters for outliers Deal with null values. Standardize variables. Address class imbalance. Train & test split, SMOTE
Data Challenges	Target class imbalance, identifying column features such as ambiguous features, limited labeled data
Modeling Techniques	Classification: - Logistic Regression - Decision Trees with ensemble methods
Target Variable	Fraud (1) or no fraud (0)
Regression or Classification problem	Classification
Tools/Methodologies	Pandas, SkLearn, NumPy

Status Update

Provide update/reminder on the goal of the project for the client here

What we did in last two weeks

Capstone Project Check-ins

The team has been meeting regularly to communicate updates and report progress

Create Presentation Outline

Create capstone PowerPoint presentation and rough outline of slides

Finalize EDA Techniques

The majority of EDA was completed, and further discussion of ML techniques was held

Discussion of Visuals and Presentation

The team discussed relevant visuals to include in the final presentation and responsibilities per team members during presentations

What we will accomplish in next two weeks

Complete all ML and Model Building

Finalize model and make last minute changes

Continue Individual EDA

Complete GitHub ReadMe file, finalize PowerPoint presentation

Presentation Practice

Perform practice runs of the final presentation

Roles and Responsibilities

Pengwei Wang: ML building and practice presentation

Hadley Campbell: ML building, practice presentation

Brian Rodriguez : ML Building, practice presentation

Payton Stauble: ReadMe file, PPT prep, practice presentation