**Milestone 1: Group Information and Communication/Work Plan (Week 1)**

Balancing work and full-time classes leaves me with little time to dedicate to group work during the week. As a result, I've opted to work on the project independently during my late nights and weekends. However, I aim on getting the peer review done as early in the week as possible so I can provide prompt feedback to my peers.

**Milestone 2 – Data Selection and Project Proposal (Week 2)**

I'm interested in working on a project focused on predicting a health ailment. The type of ailment heavily relies on the availability of high-quality public datasets. I am particularly interested in exploring diabetes datasets. By Week 2, I hope to gain more clarity on the project's direction.

Model Selection: In my last class, I built Logistic Regression, RandomForest Classifier, and DecisionTree Classifier models. I would like to explore other models like KNN or Linear Regression, but it would depend on the features in the dataset.

Model Evaluation: To evaluate my model, I intend to utilize both the classification\_report and accuracy\_score. The classification\_report will provide me with valuable information such as Precision, Recall, and F1 score, which will come in handy if my dataset is not evenly distributed.

Learnings: I am interested in delving into new models and comprehending their outcomes. Additionally, I aim to assess the potential of a model predicting the onset of a health ailment.

Risks Or Ethical Concerns: My primary concern is the genuineness of the dataset source and its legality. As I continue to reach each milestone, I will identify and include any additional ethical concerns or potential risks.

Contingency Plan: During my previous class, I encountered an issue where I began to doubt the accuracy of my models and suspected that the dataset lacked essential features to make accurate predictions. In preparation for this scenario, I started working on a similar dataset within the same field. Similarly, for my current project, I will have a contingency plan based on the same grounds.

Important Notes: Feature selection is key for model building. I plan on spending some time understanding the dataset to ensure the right features are selected for model building.

**Milestone 3 – Preliminary Analysis (Week 6)**

* Will I be able to answer the questions I want to answer with the data I have?

Based on my previous class experience, I am confident that utilizing appropriate datasets and techniques can lead to accurate answers to the right questions.

* What visualizations are especially useful for explaining my data?

I am interested in utilizing various visualization tools such as bar charts, histograms, pie charts, and scatter plots to represent data. Depending on the nature of the data, I am also open to exploring other chart options, including line charts, heatmaps, and confusion matrices.

* Do I need to adjust the data and/or driving questions?

To create accurate plots and models, it may be necessary to make modifications to the data to ensure the features are meaningful.

* Do I need to adjust my model/evaluation choices?

Utilizing StandardScalar, SMOTE, and other techniques may be necessary to modify the model evaluation methods.

* Are my original expectations still reasonable?

My goal is to meet the initial expectations while remaining open to the possibility of making small adjustments if necessary.

**Milestone 4 – Finalizing Your Results (Week 8)**

* Explain your process for prepping the data

In the modeling process, data preparation is carried out in multiple layers with the aim of completing most of the tasks on the first attempt. However, depending on the model building and outcome, there may be a need for additional preparation steps. Some of the tasks involved in data preparation include null checks, filling in missing data, dropping duplicates, renaming features, and dropping features.

As we move forward with the project, we may consider incorporating additional steps into the data preparation, modeling, and evaluation processes. I intend to keep these steps updated as well as any findings we come across during this phase.

**Milestone 5 – Final Project Paper and Presentation (Week 10)**

My plan for this phase is to present my findings through both a PowerPoint and a video presentation. The main goal is to effectively convey the outcome to the intended audience.

**Peer Review:**

For the peer review step, I would like to review the work of the following classmates.

1. Rahul Rajeev
2. Paul Galvez
3. Sara Warady