

Project Milestone One – Individual Project

Logan Quandt

College of Science and Technology, Bellevue University

DSC630 – Predictive Analytics

Dr. Fadi Alsaleem

June 12<sup>th</sup>, 2022

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

The first milestone requires that we determine whether we will complete the project in a group or individually in addition to laying out a plan for completing the following milestones. Due to scheduling constraints and limited availability, I felt that it is best if I complete the project individually. While this requires me to miss out on an opportunity to build experience working with a team on a data science project, I feel confident that I can take on the full responsibility of a project myself. I also believe this allows me to further my knowledge by being able to tackle every challenge myself.

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

As the second milestone requires us to solidify our data selection and proposal, I will spend the beginning of week 2 reviewing Kaggle, UCI's ML repository and other sources for an adequate dataset to use on the project. After the data is selected, I will review it to determine the question I would like answered from the data and answer the following questions in my project proposal:

- What types of model or models do you plan to use and why?
- How do you plan to evaluate your results?
- What do you hope to learn?
- Assess any risks with your proposal
- Identify a contingency plan if your original project plan does not work out
- Include anything else you believe is important

I will also choose a backup dataset for use as a contingency plan during this stage for use if my original project plan does not work out (ex: The models do not produce the expected results).

## **Milestone 3 – Preliminary Analysis (Week 4-6):**

I plan to start Milestone 3 in Week 4 after completing the Milestone 2 peer review in Week 3. During the third milestone I will be completing exploratory data analysis to get a better understanding of my data. This stage will also allow me to review my data to ensure I can answer the original question from milestone 2 and determine if there are any changes that are needed to be made to my model choice, data, questions, or if my preliminary expectations were reasonable. If the data is not able to answer my questions or if the data is of poor quality, I will revert to the backup dataset identified during milestone 2.

## **Milestone 4 – Finalizing Your Results (Week 7-8):**

The fourth milestone will be where I complete most of the technical work on this project. I will begin work on this during Week 7 after completing the Milestone 3 peer review. During this phase, I will describe and complete a data preparation step. (Ex: feature creation, removing null values, hiding unneeded features, etc.). After the data is prepped, I will create a model and

evaluate its performance. At this stage I will also begin to document my initial conclusions from the project.

### **Milestone 5 – Final Project Paper and Presentation (Week 9-10):**

Using the results from the model in Milestone 4, I will finalize the project including a written paper (introduction, summary of the method and results and a conclusion) during Weeks 9 & 10. In addition, I will create an 10-15 minute audio/visual presentation of slides summarizing my project.

### **Identify another project (group or individual) to perform your Milestone 2 Peer Review:**

Alexa Wittlieff and I have agreed to review each other's projects for peer reviews.