School of Business
OPIM 5604 Fall 2019 – Predictive Modeling
Project Requirements

Your task is to apply your learnings from this course to fully execute a predictive model on the dataset of your choice. Your project must follow the SEMMA process to yield valid business findings. These could be data from a problem from your current job, something of interest to the school, data acquired from the web, etc. You also will research existing solutions to the problem, and if any have been proposed or documented (on the web or elsewhere), please refer to them in your whitepaper and the presentation, articulating how similar or different the existing solution is to your team's approach/solution.

Project Proposal:

Each group has to obtain the instructor's permission regarding the appropriateness of the topic chosen. Email the instructor a project proposal with the following information about your proposed project:

- What data you have chosen, including a link to where you found the data.
- How many rows and how many columns are in the dataset.
- What you plan to predict from the data? and what might be the target variable?
- What is the exact business problem?
- What business value there might be from your findings and prediction.

Deliverable #1: On September 30th, you will share a proposal for your project. This should give as much detail as possible your ideas, so that I can give you feedback.

Deliverable #2: On October 28th, you will share with me a status report, including preliminary results or issues that you are facing in developing your project. The expectation for this report is not to deliver a polished work product but to give a reasonable update on the status of the project. This can be a paragraph or two.

Deliverable #3: On December 2 you will submit your final whitepaper, which should include the information detailed on the next page, in approximately the order given.

Presentation:

Presentations are to be no more than 10 minutes in length, followed by a brief question & answer session. All team members must speak for part of the presentation. This should be considered as an "Executive Presentation" with a significant focus on your findings, recommendations, business case for deployment and your plan for deployment. Please keep in mind that this is a very important skill to master: if a senior executive /board member tells you she'll give you 5 minutes to present your idea or proposal, you present your proposal in 5 minutes -- not 7 or 10. Going over the time will not be an option. Going over your time by more than one minute will be reflected negatively in your grade, but I'll warn you when you're getting close.

In addition to the focus areas mentioned above, he other parts of the presentation may include the following high-level themes (described in more detail in the Whitepaper section below) - 1. Business Understanding 2. Data Understanding 3. Data Preparation 4. Modeling 5. Model Evaluation.

Whitepaper:

Each team is required to create a white paper for your project. The white paper must contain the following sections listed below. **The paper cannot exceed 20 double spaced pages for the main body of the paper (excluding the title page, table of contents, references, and appendix).** Make sure your paper is readable with text and tables explaining your models and results. Don't just include a series of screen shots. Some screen shots when carefully chosen can enhance your paper, but do not rely on these exclusively to tell your story.

- Title Page
- Table of Contents
- Executive Summary – Describe at a high level what the project is about and what you concluded. If someone reads only this part of your paper, they should have a good idea of what you accomplished in your project and why it is of value.
- Business Understanding (take this seriously) identifying and defining the business problem your predictive model is potentially going to solve.
- Methodology – You may want to include sub-sections for each step in the SEMMA process. Please include details on all models that you tried, not just your best one. What is your final model and why did you choose it?
- Evaluation - Discuss how the result should be evaluated. How should a business case be developed to project expected improvement? ROI? If this is impossible/very difficult, explain why and identify any viable alternatives.
- Deployment - Discuss how the result of the data mining will be deployed. Discuss any issues the firm should be aware of regarding deployment. Are there important ethical considerations? Identify the risks associated with your proposed plan and how you would mitigate them.
- References – In order to better understand your data and the industry that it comes from, some external research may be required. Keep track of your sources and include them here. Also, your textbook and the source of your data should be included.
- Appendix

Grading:

Your group project grade will be based on the white paper (50%) and the presentation made to the class (50%). Overall, the project represents 30% of your course grade. Peer evaluations may play a role in determining the scores awarded to individual group members on the project Individuals who assume leadership roles and/or who perform more than their fair share of the work and/or whose work is consistently of superior quality are likely to receive scores that are above the group's average scores. Individuals who do less than their fair share of the work or work that is consistently below average will receive scores that are less than the group average.