DSC 630 Course Project

Course Project Schedule

Week 1: Milestone 1 Due (Team Information/Communication Plan)

Week 2: Milestone 2 Due (Data Selection and Project Proposal)

Week 3: Milestone 2 Group Evaluation or Peer Review

Week 6: Milestone 3 Due (Preliminary Analysis)

Week 7: Milestone 3 Group Evaluation or Peer Review

Week 9: Milestone 4 Due (Finalizing Your Results)

Week 10: Milestone 4 Group Evaluation or Peer Review

Week 12: Milestone 5: Final Project Submission (Final Paper and Presentation – Due Saturday!)

Project Overview

You can choose to do your project as a group with up to three team members or individually. For those working in teams, it is expected that all team members will contribute equally and that everyone will take the opportunity to learn from each other. If you are working individually, you will be required to do three peer reviews of other project milestones. Those working in a group will have the option to do peer reviews of project milestones or do an evaluation of your group members.

Students will identify a business problem to address through predictive analytics. Students will formulate the problem, identify the right sources of data, analyze data, and prescribe actions to make decisions on relevant questions. The goal is to select appropriate models and apply the respective methods to enhance data-driven decision making related to the business problem.

Here are some online resources where you might look for data.

- Kaggle
- UCI Machine Learning Repository
- Nature Scientific Data
- US EPA
- Human Development Reports
- National Center for Health Statistics
- US Census Bureau (econ, population, geographic, health data)
- Bureau of Justice Statistics
- Statistics Resources Online
- USA Federal Government Data
- Google Public Datasets

There are no restrictions on what data you choose, except you cannot use a data source used for course assignments. You will want to choose a dataset that is large and complex enough so that you can demonstrate the use predictive analytics concepts and methods and make a nontrivial conclusion/recommendation based on your analysis. You are free to use Python or R (or a combination of both) for your project. There are four milestones that will be due throughout the term to help keep you on track with the project. You will submit your final project at the end of the term in the form of a paper and an audio/video presentation.

At minimum, your final project paper should include the following.

Introduction

- Problem statement
- Explain why the problem is important/interesting
- Who would be interested in solving this problem, i.e., who would you be trying to sell this project to?
- Where did you get your data?
- Why is this data useful to solve the problem?

Methods/Results

- What did you find out by exploring the data?
- Are there any visualizations that help tell a story with your data?
- What steps did you perform to prepare the data?
- What type of types of modeling are you using on your data?
- What metric(s) are you using to measure your results?
- Why did you choose the metric(s) you chose?

Conclusion

- What did you learn?
- What recommendations would you make based off your analysis?
- Is your model ready for deployment?
- What work still needs to be done?

References

- Include at least three properly cited references at the end of your paper
- Also include in-text citations

Milestone Peer Reviews or Group Evaluation

In Week 3, Week 7, and Week 10, you will be required to submit milestone peer review or a group evaluation.

If you are working in a group, you have the option to submit a peer review of another project (group or individual project) or submit a group evaluation of your own group. There are templates to fill out for each of these options. If you are working individually, you will have to complete a peer review of another project (group or individual).

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

The first step is to determine if you are doing the project independently or as a group. Groups can have a maximum of three students. For each project group, whether you are working individually or as a group, create a folder (name it so that your group is easily identifiable) in the class Teams channel where you will put your project milestones for peer reviews. You will also submit your milestones in Blackboard for grading. Based on whether you are working as a group or independently, follow the below instructions for the submission.

Group Project Requirements

- Submit team member names
- Determine your method of communication as a team (Slack, Teams, Zoom, etc.)
- Identify a time that you can meet on a weekly basis to discuss progress and next steps
- Provide high level plan for tackling the project

Independent Project Requirements

- Craft a rough project plan for how you will accomplish each milestone and task, what your project work schedule will look like, etc.
- Identify another project (group or individual) to perform your Milestone 2 Peer Review

Submit the requirements to the assignment link provided. This information will be used for the instructor to create a group within Blackboard that will have a location for assignment submission. If you are working independently, you will still have a group created for you, but will only be visible by you and the instructor. Milestone 1 is the only one that is not submitted via your group.

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

Data selection and your project proposal are due this week. While you might decide to add additional data sources as the project progresses, you should have a good idea of your initial dataset by this milestone.

Milestone 2 should include the information outlined in the introduction above. Additional items to address are the following.

- 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

The proposal should be a minimum of three pages, double-spaced. You should treat this proposal as the start of your final project paper submission. But also remember this is only the initial proposal – your findings might take you in a different direction for the final submission.

Please submit Milestone 2 in Blackboard under the group submission link. Also, post your Milestone 2 in your Teams project folder for peer reviews.

Milestone 3 – Preliminary Analysis (Week 6)

Milestone 3 should include all the information from Milestone 2, updated if necessary. Additionally, you should start to explore your data. Items to address include the following.

- Will I be able to answer the questions I want to answer with the data I have?
- What visualizations are especially useful for explaining my data?
- Do I need to adjust the data and/or driving questions?
- Do I need to adjust my model/evaluation choices?
- Are my original expectations still reasonable?

Please submit Milestone 3 in Blackboard under the group submission link. Also, post your Milestone 3 in your Teams project folder for peer reviews.

Milestone 4 – Finalizing Your Results (Week 9)

In Milestone 4, most of the technical work for the project should be done. You should include the information from Milestone 3 and address the following additional items.

- Explain your process for prepping the data
- Build and evaluate at least one model
- Interpret your results
- Begin to formulate a conclusion/recommendations

Please submit Milestone 4 in Blackboard under the group submission link. Also, post your Milestone 4 in your Teams project folder for peer reviews.

Milestone 5 – Final Project Paper and Presentation (Week 12)

Note that you need to submit two items for your final project submission. The paper should include an introduction, a summary of your methods and results, and a conclusion as outlined above. In addition, submit an audio/video presentation with slides summarizing your project. A good goal for the length of your presentation is 10-15 minutes. Think about this as a high-level presentation you would give to your CEO.