

MSIS 5633 – Predictive Analytics Technologies

Group Term Project

Term Project Description:

One of the most important and most value adding activities of this course is the Group Term Project, where you will get to analyze a real-world situation using a large and feature rich real-world database by leveraging the capabilities of one of the most popular data mining tools—KNIME. Some of the major project activities that you will perform include downloading and merging the raw data, properly preprocessing/cleaning/transforming the data (e.g., removing unrelated rows and column, aggregating certain variables, handling missing values, etc.), developing several prediction models, assessing the viability and accuracy of these models, and interpreting/discussing the variable importance scores. To integrate and/or preprocess the data, you can use any software that you are familiar with, however, for a comprehensive exposure to the software platform, I recommend to you to use KNIME for everything.

Term Project Deliverables:

1. A Final Report and a PPT File

- Put together a professionally organized document that includes a cover page, an executive summary, and subsections that follow the six-step CRISP-DM process model. Prepare the document in MS Word, and then convert it to a single pdf file.
- The length of the document should be 15-25 pages, using 12-point times new roman fonts, 1” margins, and 1.5 line spacing.
- Make sure to embed appropriate screen shots, other figures/charts and tables where they belong within the document to maintain logical flow and continuity. When including screen shots, only include the section of the screen that matters, making sure they are crisp and readable with good enough resolution.
- Submit both files (pdf document and the PowerPoint file) to course site by the due date/time.

2. A Presentation

- You will get to make a 20-25 min team-based presentation to summarize your project activities and discuss your findings. During the presentation, because of the time limit, you can use a somewhat shorter version of your presentation slide deck.

Due Dates/Times:

- Final Report & PPT File: May 1, 2022, at 11:55 PM.
- Team Presentation: April 27, 2022, at 11:55 PM.