Team Project Requirements and Guidelines

Project Objective:

Able to research and explain a real-world application of AI/ML in Business, Government or Social Good to management (non-expert) audience. i.e. Your team can choose a project of interest.

Project Deliverables:

1. Project Report in Word Document Format.
2. Datasets, preferably in CSV format, or provide link to datasets if size is too big to upload.
3. Analysis Scripts (Rscript/Python script/SAS/Excel files)
4. Presentation Slides in PowerPoint format.

Project Report General Requirements and Guidelines:

1. Format: Microsoft Word document. Max 20 pages (excl. appendix and executive summary).
2. One page executive summary that summarize most important findings.
3. Clearly define the Business problem (or opportunity) statement(s).
4. Explain the business outcome measures and desired targets (if any).
5. Explain the AI/ML solution and how it solved the business problem or opportunity.
6. Define and explain the AI/ML performance measures and targets.
7. Find suitable dataset(s) that can be used to explain at least one component of the AI/ML solution.
8. Demonstrate and explain at least two permitted techniques using a software (R/Python/Excel). The script will need to be submitted too. The submitted dataset and analysis files/scripts must be sufficient for someone to reproduce and verify your work done.
9. Refer to Appendix A for Exclusions and Appendix B for Permitted Techniques.
10. The entire AI/ML solution (if multiple components) must be explained. The dataset and demonstration can choose to focus on just one component of the solution due to limitations of time and access to actual data.

Presentation Slides General Requirements and Guidelines:

1. A deck of PowerPoint presentation slides in PPTX format that summarize your important findings and recommendations.
2. All students must speak and present their slides.
3. State the slide speaker name in a corner of each slide so that marks can be attributed to that speaker.
4. Duration: max 25 minutes presentation followed by 5 mins Q&A.

- The presentation is during class time in the last session (assuming class duration is sufficient). Your team will be randomly assigned a presentation timeslot unless you have good reason for a specific timeslot.

Popular websites to search and download Datasets:

- https://www.kaggle.com/datasets
- https://archive.ics.uci.edu/datasets
- https://data.gov/
- https://beta.data.gov.sg/
- Many other data sources …

Target Audience:

The target audience for the report and presentation is senior management who are typically not familiar with AI/Machine Learning. You may include the more technical details in appendix. The focus is on the business problem/opportunity and how the business problem/opportunity was/can be solved using AI/Machine Learning.

Submission:

Put all the deliverables and supporting files/materials/datasets in a zipped file without password and submit your zipped file into NTULearn > Team > File Exchange. Include your class and team number in all the file names, including the zipped folder. If the dataset is too big, provide a download link.

List the names of all your project team members on the first page of the project report and the first slide in presentation slides.

The submitted file serves as the record of your submission time. If you need to make any corrections or edits after the deadline, you can submit a revised version within 24 hours of your team presentation time slot but do not delete the previously submitted file. Name those file(s) or new zipped folder with the word "REVISED".

Failure to comply with the instructions listed in this document may result in marks penalty.

Submission Deadline:

Refer to latest Course Schedule in NTULearn > Information.

IMPORTANT: If any team-mate did not contribute sufficiently to Team Assignment and Team Project, please inform your class instructor and submit peer evaluation by last day of the course so that marks can be downgraded.

If all your team-mates contributed sufficiently, then you do not need to submit peer evaluation.

Appendix A: Exclusion List

Do not use any of the below organization unless approved by instructor for exceptional reasons.

1. Airbnb
2. Amazon.com
3. All Bike sharing companies
4. All ride-sharing companies (e.g. Grab, Uber, Gojek, …etc)
5. Netflix
6. Spotify
7. Target
8. Walmart
9. Any project/research/analysis/report already submitted or plan to be submitted for another course. Your project submission must be your team's original work and not submitted to any other course.

Appendix B: Permitted Techniques List

You must use at least two techniques in category A and may choose to use techniques in category B.

Seek instructor approval if you want to use other technique(s) not in the lists.

The grading will mainly be based on techniques learnt in this course. Other techniques outside this course (if any) are to be used for reference/comparison purpose.

Category A:

1. Linear Regression
2. Logistic Regression
3. CART
4. Facet Chart [see ggplot solution in Unit 4].

Category B:

1. Any Descriptive Statistics taught in first course in Statistics.
2. Any Inferential Statistics taught in first course in Statistics.
3. Clustering
4. Lasso Regression
5. Ridge Regression
6. Text Mining
7. Bootstrap
8. Time Series Forecasting
9. Random Forest
10. XGBoost