

Assignment-1 sneak pick

In this assignment you will explore a modified real dataset and practice the typical machine learning process. This assignment is designed to help you become more confident in applying machine learning approaches to solving tasks.

To complete this assignment, you will require skills and knowledge from lecture and lab material for Weeks 1 to 3 (inclusive). You may find that you will be unable to complete some of the activities until you have completed the relevant lab work. However, you will be able to commence work on some sections as early as the end of this week.

Thus, do the work you can initially, and continue to build in new features as you learn the relevant skills. A machine learning model cannot be developed within a day or two. Therefore, start early.

This assignment has three deliverables:

1. A bullet point style PDF report, preferably in the form of notebook (this will be explained in your next lecture). Regardless, your report should include the graphs produced by your analysis. If you are using notebook, it needs to be in the format of the provided tutorials. That means the report should include markdown text explaining the rational, critical analysis of your approach and ultimate judgement. The report needs to be self-explanatory, well structured, and fulfill all the assignment specifications.
2. A video presentation that goes along with your PDF report explaining your steps in sufficient detail and with proper judgment (this will be explained in your next lecture).
3. Your Python scripts or Jupyter notebooks used to perform your modelling & analysis with instructions on how to run them, which need to have embedded explanatory comments.

More detail of the assignment is below:

- Weight: 30% of the final course mark
- Type: Individual Due Date: 8th of April 2024 (Week 5)
- Learning Outcomes: This assignment contributes to CLOs: 1, 3, 4

To complete this assignment successfully, you need to know, how to properly perform bellow steps for the given tasks, using python:

- **Load data**
 - data will be provided in excel format; you will be also given meta data to provide you with more insight about your data (i.e. explanation about what each column / input feature represents).
- **Start your EDA**
- **Data normalization (no feature selection at this stage)**
- **Examine features and data**
- **Setup evaluation framework**
- **Try initial model (e.g. linear regression, etc.)**
- **Evaluate and update parameters and/or models (e.g. using polynomial regression, etc.)**