Coursework 1 (group) – ML end-to-end Project

In this group assessment you will be provided with a data set to use as your input for an end-to-end ML project. The aim is to go through all the stages of the ML project making decision which you can justify. The ultimate goal is to answer a particular question for the given case through ML, and develop insights into the specific case. Finally you will present your approach and findings to your colleagues in class.

These are the elements for the mark scheme:

- Data preparation - 20%
    - Preliminary analysis of the provided data
    - Ensure data is of good quality and usable
    - Establish strategy for data preparation
    - Prepare the data for the necessary use
- Data analysis and visualization -20%
    - Analysis of prepped data
    - Provide visualization of relevant aspects for ML model selection
    - Establish ML strategy
- Model selection and training – 20%
    - Select ML model
    - Implement and train solution
    - Demonstrate steps and iterations
- Results and analysis – 20%
    - Visualization of results
    - Performance analysis of results
    - Implications of solution and critical analysis
- Presentation skills – 20%
    - Clarity of presentation
    - Group participation
    - Conclusion analysis

CW1 Cases

Case 1 – Climate change data

- Establish a prediction model for the UK for temperature
- Compare model effectiveness for similar countries

Case 2 – Sports performance data for Powerlifting prediction - Powerlifting Performance

- Establish a prediction model for total lifting weight based on age of an athlete for the most popular weight classes
- Compare effectiveness of model for male and female.

Case 3 – Lung cancer data

- Establish a classification model for lung cancer prediction and assess its performance metrics
- Establish critical features correlated to lung cancel

Case 4 – Salary data

- Establish a classification model for salary prediction and assess its performance metrics
- Establish critical features correlated to higher salary

Case 5 – Product demand data

- Establish a prediction model for product with highest demand
- Compare effectiveness of model for demand of “Product_1295”

Case 6 - Failure Prediction

- Establish a classification for the different types of failures
- Compare the different features that correlate with the “Overstrain Failure”