Coursework 2 (Individual) – ML end-to-end Project

In this individual assessment you are required to select an appropriate data set from https://www.kaggle.com/datasets, which should fit to your interests. The selection of the data set should consider its appropriateness of the data for the ML project. The student will need to formulate the ML objective, so also take this into account when looking at the data sets.

Once this is done the student is expected to through all the stages of the ML project and add commentary justifying each step. The final step is to comment on the outcome of the project in the context of ML challenges and limitations covered in class.

Submission: Jupyter Notebook File and document with ML end-to-end project commentary.

Please note: Commentary on code is helpful and will be reflected in marks

These are the elements for the mark scheme:

- Data set selection and definition of ML objective 10%
    - Data set explanation and justification for selection
    - ML objective and motivation
- Data preparation and Analysis - 25%
    - Ensure data is of good quality and usable
    - Establish strategy for data preparation
    - Prepare the data for the necessary use
    - Analysis of prepped data
    - Provide visualization of relevant aspects for ML model selection
- Model selection and training – 25%
    - Analyse potential ML models
    - 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
- ML Challenges and limitations – 20%
    - Identification of issues with the case
    - Propose mitigation measures
    - Case conclusions summary