Compulsory Assignment #1

Machine Learning and Deep Learning (CDSCO2041C)

Somnath Mazumdar
sma.digi@cbs.dk
Department of Digitalization, Copenhagen Business School

Deadline: Check Canvas LMS

Instructions

1. Write your group members' student IDs in the report.
2. If you have additional content/text to present, you may include it in the notebook itself.
3. The complete solution code must be submitted as **one single *Jupyter* notebook**.

Assignments

Question 1

Exploratory Data Analysis (EDA)

The provided dataset contains UK territorial greenhouse gas emissions by source and activity, covering the period from 1990 onwards. Emissions are attributed to the sector that emits them directly and include indicators related to UK territorial totals, international aviation and shipping, and Paris Agreement coverage.

Perform Exploratory Data Analysis (EDA) to investigate the key factors driving changes in UK greenhouse gas emissions over time.

1. Write a Python program to perform a covariance- and correlation-based analysis to examine relationships between emissions, sources, and activities across years. **Do not** use any built-in covariance or correlation functions. You must implement your **own** calculations.
2. Write another Python program to visualise your findings from the previous step and briefly explain the observed emission patterns. *Hints: You may consider using* histograms*,* boxplots*, and* scatterplots.

Question 2

Cluster Analysis

Cluster analysis is used to group data points based on similarity in their attributes.

1. Choose one clustering algorithm discussed in the lectures and apply it to group emission sources or activities based on their emission trajectories over time.
2. Clearly justify the variables used for clustering (e.g., emission levels, rate of change, fuel group). Interpret the resulting clusters and explain what they reveal about structural changes in UK greenhouse gas emissions.
3. Relate your clustering results to UK climate policy by identifying which clusters align with sectors targeted under UK Carbon Budgets and which sectors appear more resistant to emission reductions.

Question 3

Policy Interpretation and Critical Analysis

1. Using the column `Included in UK territorial total`, compare emission trends with and without international aviation and shipping. Discuss how this distinction affects progress assessment against UK Carbon Budgets.
2. Using the `Included in UK Paris Agreement total` indicator, identify which emission sources are covered under the UK's Paris Agreement reporting. Explain the implications of this coverage for interpreting national emission reduction performance.
3. Based on your data-driven findings, critically assess whether historical emission trends suggest that the UK is structurally aligned with its long-term climate targets. Support your answer with quantitative evidence from your analysis.