Entering Practical Exam for the Decision Sciences Research Center at Tecnologico de Monterrey

Dear Applicant,

Thank you for applying to the Research Programmer and Data Scientist position at the Decision Sciences Research Center at Tecnologico de Monterrey. We have designed this practical exam to assess your technical skills, analytical thinking, and ability to work with real-world data. Please ensure that all your answers are thoroughly documented in a GitHub repository to ensure replicability.

Below are the questions you are required to solve. For each question, provide a script that performs the required tasks, and document your methodology, assumptions, and results in a Markdown file. A comprehensive README file should be included in your GitHub repository, explaining how to set up the environment, run your scripts, and interpret the outputs.

Question 1: Comprehensive Data Acquisition and Preprocessing

Task:
Download and preprocess CO2 emissions data along with a wide range of socio-economic and environmental indicators from the World Bank's Climate Change database.

Instructions:
1. Access the World Bank database using Python, R, or MATLAB.
2. Download CO2 emissions data and as many relevant socio-economic and environmental indicators as possible (e.g., GDP, population, energy consumption, urbanization rate, education level, etc.).
3. Clean and preprocess the data, addressing missing values, outliers, and ensuring consistency across indicators.
4. Provide a detailed summary of the dataset, including key statistics, correlations between variables, and any notable patterns or anomalies.

Deliverable:
- A script that downloads, preprocesses, and summarizes the data with an emphasis on integrating a wide range of predictors.
- A Markdown file summarizing key statistics, correlations, and preprocessing steps.

Question 2: Predictive Modeling and Scenario Analysis

Task:
Develop a predictive model to forecast CO2 emissions based on the comprehensive set of indicators. Use the model to answer an analytical question: "If a country increases its GDP by 10%, what is the expected percentage change in CO2 emissions, assuming all other factors remain constant?"

Instructions:
1. Split the data into training and testing sets.
2. Train a regression model using the socio-economic and environmental indicators as predictors.
3. Evaluate the model's performance using appropriate metrics (e.g., RMSE, R^2).
4. Simulate the scenario where GDP increases by 10% across different countries, keeping other factors constant.
5. Analyze and interpret the results, providing a range of expected changes in CO2 emissions across nations.

Deliverable:
- A script that trains the predictive model, evaluates its performance, and simulates the GDP scenario.
- A Markdown file explaining the choice of model, scenario analysis, and interpretation of the results.

Question 3: Fermi Problem and Sensitivity Analysis

Task:
Estimate the impact on global CO2 emissions if 50% of the world's population adopted electric vehicles (EVs). Use your model to perform sensitivity analysis and answer: "Which countries would see the most significant reduction in emissions, and by how much?"

Instructions:
1. Integrate relevant indicators such as the number of vehicles per capita, energy mix, and current EV adoption rates into your model.
2. Perform a sensitivity analysis by simulating a scenario where 50% of the population in each country adopts EVs.
3. Identify the countries with the most significant reductions in CO2 emissions and quantify the impact.
4. Discuss the assumptions made in this estimation and their potential limitations.

Deliverable:
- A script that performs the sensitivity analysis and estimates the impact of EV adoption.
- A Markdown file discussing the findings, key drivers of emission reduction, and assumptions.

Question 4: Classification and Policy Implications

Task:
Build a classifier to identify countries that are likely to achieve a significant reduction in CO2 emissions in the next decade. Answer the business case question: “What are the common characteristics of countries that successfully reduce emissions, and how can policymakers in other nations apply these insights?”

Instructions:
1. Define a binary target variable indicating whether a country is likely to reduce its CO2 emissions significantly in the next decade.
2. Train a classification model using the comprehensive set of indicators.
3. Evaluate the classifier's performance using metrics like accuracy, precision, recall, and F1 score.
4. Analyze the key features driving the classification and identify common characteristics of successful countries.
5. Provide policy recommendations based on the insights gained.

Deliverable:
- A script that builds and evaluates the classifier.
- A Markdown file interpreting the classifier's performance, key features, and policy implications.

Question 5: Strategic Analysis and Model Application

Task:
Using the predictive model and classifier, conduct a strategic analysis to answer: “If a country were to invest heavily in renewable energy, what is the likelihood that this investment would lead to a reduction in CO2 emissions within the next five years? How should this country prioritize its investments to maximize impact?”

Instructions:
1. Use the predictive model to simulate the impact of increased investment in renewable energy on CO2 emissions.
2. Combine the results with the classifier to assess the likelihood of emissions reduction.
3. Prioritize different types of investments (e.g., solar, wind, hydro) and regions within the country to maximize the impact.
4. Provide a strategic recommendation report that includes a ranked list of investment priorities and the expected outcomes.

Deliverable:
- A script that performs the strategic analysis and simulates investment scenarios.
- A Markdown file with a strategic recommendation report, including the prioritization and expected outcomes.

Please submit your completed exam by sharing the link to your GitHub repository containing all scripts, Markdown files, and the README file. Ensure that your work is well-documented and that the repository is accessible.

We look forward to reviewing your submission.

Best regards,
The Hiring Committee
Decision Sciences Research Center
Tecnologico de Monterrey