+NUTRI Co

CASE CHALLENGE

INSTRUCTIONS

- This challenge is designed to evaluate the potential of candidates for the role of Data Scientist.
- Candidates are expected to demonstrate their skills and reasoning abilities through a series of questions covering different areas of data science.
- The dataset provided contains detailed nutritional information for a wide range of food items.
- Candidates are encouraged to approach the questions with creativity and demonstrate their ability to derive insights and build predictive models using the dataset.

QUESTION 1: EDA (5 POINTS)

- Using the provided dataset, perform an exploratory data analysis that reveals insights into the nutritional content of food categories.
- Identify patterns, outliers, and any interesting correlations between different nutrients.
- Present your findings with clear visualizations and a brief commentary on the implications of your analysis.

QUESTION 2: ML (5 POINTS)

- Build a machine learning model that predicts the fat content of a food item based on its other nutritional information.
- Explain your choice of model, feature selection, and any data preprocessing steps.
- Evaluate the model's performance using appropriate metrics and discuss how the model could be improved or applied in a practical context.

QUESTION 3: DL (5 POINTS)

- Propose and implement a deep learning approach to classify food items into their respective categories based on their nutritional information.
- Detail your neural network architecture, training process, and how you address challenges such as overfitting.
- Assess the model's accuracy and discuss potential use cases for such a classification system in real-world applications.

+NUTRI Co

CASE CHALLENGE

QUESTION 4: ARCHITECTURE (5 POINTS)

- Considering the results obtained from questions 1-3, suggest a data architecture or platform that would allow users to interact with these results.
- Your proposal should include how users will access the models' insights, how the data should be stored, and how the models should be served and updated.
- Be creative and justify your choices based on the dataset's context and potential use cases.

GRADING

Each question will be graded on a scale of 1 to 5, with 5 being the highest. The grading will consider the correctness of the solution, creativity, clarity of presentation, and depth of analysis.

- Correctness and Completeness (2 points): The solution correctly addresses the question with comprehensive analysis.
- Creativity and Innovation (1 point): The solution demonstrates creativity in approach and analysis.
- Clarity and Presentation (1 point): Insights and solutions are clearly presented, with well-designed visualizations for data-driven points.
- Depth of Analysis (1 point): The solution goes beyond surface-level analysis, offering deeper insights or identifying underlying patterns.

Submissions should include both a PowerPoint presentation and the source code used for analysis and modeling. The candidate will present their solution during an interview, where they will be evaluated on their ability to articulate their thought process and findings.
Solutions should be sent to **sergio@nutrico.io** within a maximum of 3 days after reception of the assignment.

GOOD LUCK!