WARDLIN

WardLin
352 55 Växjö
kronoberg , Sweden
www.wardlin.com
Tel:+46700265148

Test Assignment for Data Scientist position at WardLin:

Dear Candidate,
Thank you for your interest in the Data Scientist position at WardLin. We are excited to have you participate in this stage of our selection process.
This test is designed to evaluate your data analysis skills, statistical knowledge, and problem-solving capabilities, which are critical for the role. Please read the instructions carefully and ensure that you complete all sections to the best of your ability.

Task Description:

1. Create a synthetic dataset representing environmental parameters and corresponding impact scores. You can use any method or library to generate this dataset, ensuring it reflects realistic scenarios.
2. Perform exploratory data analysis (EDA) to understand the dataset's characteristics and relationships between features.
3. Pre-process the data as needed, including handling missing values, encoding categorical variables, and scaling numerical features.
4. Split the dataset into training and testing sets.
5. Develop a predictive model to estimate the environmental impact score based on the given features. You can choose any suitable machine learning algorithm(s) for this task.
6. Evaluate the performance of your model using appropriate metrics.
7. Provide insights into which features contribute the most to the environmental impact assessment.
8. Perform data cleaning and transformation as necessary to prepare the data for visualization.
9. Write a brief explanation of each visualization, describing the insights it provides and how it contributes to understanding the environmental impact assessment.

Deliverables:

- Jupyter Notebook or Python script containing your code, comments, and explanations.
- A brief report (PDF format) summarizing your approach, the rationale behind your model selection and visualization choices, insights gained from the analysis, and any potential areas for improvement.

Evaluation Criteria:

Your submission will be evaluated based on.

- Correctness and efficiency of the implemented solution for predictive modelling.
- Clarity and organization of the code and comments.
- Depth of analysis, insights, and interpretation provided in the report.

Notes:

- Feel free to reach out if you have any questions or need clarification on the task requirements.
- We value creativity and innovative approaches in solving problems and visualizing data, so don't hesitate to showcase your skills!

Submission Instructions:

- Generate the synthetic dataset according to the provided instructions.
- Develop your solution using Python for the predictive modelling part and Power BI Desktop for data visualization.
- Submit your Jupyter Notebook/Python script and the report in PDF format via email to info@wardlin.com no later than Saturday, 25th of May, 2024, at 9:00 PM.

We wish you the best of luck