Fireai

Your trusted AI for Data analysis

Data Scientist Assignment

Objective:
This assignment is designed to assess your analytical, problem-solving, and technical skills. We are looking for candidates with a strong foundation in data science, who can demonstrate proficiency in data analysis, machine learning, and presenting data-driven insights.

Instructions:

1. **Submission Format**:
    - Submit your work in a Jupyter Notebook (Python).
    - You should include all code, comments, and explanations within the notebook.
2. **Dataset**:
    - Attached in email

Problem Statement:

You are required to perform the following tasks:

1. Data Understanding and Cleaning:

- Load and explore the dataset.
- Handle missing or incorrect data if present.
- Perform basic statistical analysis (mean, median, mode, variance) on numerical features.

2. Exploratory Data Analysis (EDA):

- Identify patterns and insights in the dataset through visualizations and summary statistics.
- Use appropriate visualizations such as histograms, box plots, scatter plots, and correlation matrices to present your findings.
- Discuss any interesting trends or correlations you observe.

3. Feature Engineering:

- Create any necessary new features (derived features) that could improve model performance.
- Provide justification for why these features might be helpful.

4. Modeling:

- Split the data into training and testing sets.
- Implement at least two machine learning models (e.g., Linear Regression, Decision Trees, Random Forests, SVM, etc.).
- Compare model performance using appropriate evaluation metrics (e.g., accuracy, precision, recall, RMSE, R-squared, etc.).
- Select the best-performing model and provide reasons for your choice.

5. Model Tuning:

- Perform hyperparameter tuning on the selected model to improve its performance.
- Discuss how hyperparameters impact model performance.

6. Interpretation & Insights:

- Provide a summary of your findings and insights from the model.
- Discuss how the model's output can be used to make business decisions.
- Mention any limitations or challenges faced during the analysis.

7. Presentation:

- Provide clear and concise explanations of each step and the rationale behind your choices.
- Include visualizations where relevant to support your explanations.

Bonus (Optional):

- Apply advanced techniques like PCA, feature selection, or ensemble methods (e.g., XGBoost).
- Work on a classification task if your dataset involves labeled categories (e.g., fraud detection, customer segmentation).

Data Analyst Assignment

Objective:
This assignment is designed to assess your ability to analyze data, generate insights, and create meaningful visualizations. We are looking for candidates with strong analytical and communication skills who can turn raw data into actionable insights.

Instructions:

- **Submission Format**:
  - Submit your work in an Excel file, Jupyter Notebook (Python with Pandas), or Power BI/Tableau (if applicable).
  - Include all relevant analyses, visualizations, and summaries.
- **Dataset**:
  - Attached in email

Problem Statement:

You are required to perform the following tasks:

1. Data Understanding and Preparation:

- Load and explore the dataset.
- Handle missing or incorrect data (if any).
- Clean the dataset by removing or imputing null values and correcting any data discrepancies.
- Perform basic statistical analysis (mean, median, mode, standard deviation) on numerical features.

2. Exploratory Data Analysis (EDA):

- Conduct an exploratory data analysis to understand trends, distributions, and correlations in the data.
- Use Python (Pandas, Matplotlib, Seaborn) oe excel formula for data visualization
- Write a summary of key findings based on your analysis.

3. Data Visualization:

- Create at least 3 relevant visualizations that provide clear insights from the data.
- Ensure that each visualization is well-labeled and clearly explains the trend or relationship it depicts.
- Include visualizations that compare different segments or groups within the dataset.

4. Insights and Recommendations:

- Based on your analysis, provide actionable insights that could help the business make informed decisions.
- Present 3 key findings and explain insights.

5. Dashboard (Optional):

- If you are using Power BI or Tableau, create an interactive dashboard summarizing your findings.
- Ensure the dashboard is user-friendly and allows easy exploration of the data.
