Heliverse

Assignment: AI & ML Role - Python Programming

Objective:

This assignment aims to evaluate your proficiency in Python programming and your understanding of key concepts in artificial intelligence (AI) and machine learning (ML).

Instructions:

You are tasked with completing the following assignment, demonstrating your skills in AI and ML using Python programming:

Problem Statement: Predicting Employee Attrition

Dataset: IBM HR Analytics Employee Attrition & Performance

Instructions:

1. Dataset Analysis and Preprocessing:
    - Download the IBM HR Analytics Employee Attrition & Performance dataset from a reputable source (e.g., Kaggle).
    - Analyze the dataset to understand its structure and features. It contains various attributes related to employee demographics, job roles, satisfaction levels, performance ratings, etc., along with a target variable indicating whether an employee has left the company (Yes or No).
    - Perform preprocessing steps such as handling missing values, encoding categorical variables, and scaling numerical features if necessary.
2. Model Development:
    - Split the dataset into training and testing sets.
    - Choose suitable machine learning algorithms (e.g., logistic regression, random forest, support vector machine) for binary classification.
    - Implement the selected algorithm(s) using Python libraries like scikit-learn or pytorch.
    - Train the model(s) on the training data and evaluate their performance using metrics such as accuracy, precision, recall, and F1-score.
3. Model Evaluation and Optimization:
    - Analyze the performance of the trained model(s) using evaluation metrics.

- Explore techniques for model optimization, such as hyperparameter tuning, feature selection, or model ensemble methods, to improve performance.
- Optimize the model parameters and evaluate the impact on model performance.

4. Documentation and Presentation:
    - Document the entire process, including dataset analysis, preprocessing steps, model development, evaluation results, and optimization techniques used.
    - Prepare a concise report summarizing your findings, insights gained, challenges encountered, and recommendations for reducing employee attrition.
    - Optionally, create a presentation to showcase your work, highlighting key steps, results, and conclusions.

Submission:

Submit your assignment as a google drive link (public) or github containing:

- Python code files (.py) implementing the data analysis, model development, and evaluation.
- A README file providing instructions for running the code and any additional information.
- A PDF report summarizing your findings and insights.
- A video presentation for showcasing your work (optional).

Note: If you encounter any technical difficulties or have questions regarding the assignment, please feel free to reach out to us at pm@heliverse.com or anmol@heliverse.com.