Student Support Team
Office Academic Lab

N1C

National CyberCity
Narawat Street and
ShweTawWin Street2

Purpose: To enhance students' logical and critical thinking skills in computer science and programming in Myanmar and globally. The Quality Control Department of National Cyber City creates each and every question.

Programming Language and Framework: **Python**
Question:
Title: Predictive Modeling with Linear Regression or any other algorithms in Python

Design and implement a Python program to demonstrate the use of the Linear Regression algorithm for predictive modeling. Your program should include the following components:

1. Dataset Selection:
   - Choose a dataset suitable for linear regression analysis. The dataset should contain relevant features and a target variable that you aim to predict.

2. Data Preprocessing:
   - Implement data preprocessing steps, including handling missing values, scaling features if necessary, and encoding categorical variables.

3. Linear Regression Model:
   - Use a linear regression algorithm to build a predictive model. You can utilize libraries such as scikit-learn for implementing the linear regression model.

4. Data Splitting:
   - Split the dataset into training and testing sets. Use a significant portion for training and reserve a separate portion for evaluating the model's performance.

5. Training the Model:
   - Train the linear regression model using the training dataset. Monitor and record relevant metrics such as the coefficients and intercept.

6. Model Evaluation:

- Evaluate the model's performance on the testing dataset. Use appropriate evaluation metrics such as Mean Squared Error (MSE), R-squared, or any other relevant metric for regression problems.

7. Visualization:
- Visualize the regression line along with the actual and predicted values on a scatter plot. Provide insights into how well the model fits the data.

8. Prediction:
- Implement functionality to accept user input for relevant features and use the trained model to make predictions. Display the predicted output along with a confidence interval if applicable.

9. Documentation:
- Provide comments and documentation for your code, explaining each step of the linear regression implementation.

10. Parameter Tuning:
- Experiment with different hyperparameters or features to observe the impact on the model's performance. Discuss any observations and insights gained.

11. Challenges and Limitations:
- Discuss any challenges faced during the implementation, and identify potential limitations or assumptions of the linear regression model.

12. Future Improvements:
- Suggest possible improvements or extensions to enhance the predictive model. This could include trying different regression algorithms, incorporating more features, or using more sophisticated techniques.

13. Submission:
- Submit the Python program along with a report detailing the dataset, methodology, results, and insights gained from the linear regression analysis.

Feel free to customize or extend the assignment based on the complexity level you desire for the course or project.

Time Consumption: This program should take roughly 10 days to complete.

Please use the following link to upload your program's file. Don't forget to zip up the entire program files. We only accept .zip file type.

Link: https://forms.gle/r7PMCqCiXQfgQdKK9
Student Support Services
National Cyber City
Narawat Street and ShweTawWin Street2
w/51 ,
+959453741976

Founder and CEO
Win Htut