Title: Analysis of Air Quality Data Using Python

Introduction:

In this report, we utilize Python programming language along with libraries such as Pandas, NumPy, and Matplotlib to analyze air quality data obtained from the AirQualityUCI dataset .This report presents the implementation of gradient descent for multivariate regression using Python. The objective is to predict target variables based on a set of features from the 'AirQualityUCI' dataset.

Code Overview:

The provided code consists of functions to compute the cost and gradient, perform gradient descent, and evaluate the model's performance. It utilizes libraries such as NumPy and Matplotlib for numerical computations and visualization.

Data Loading:

We begin by loading the dataset from the Excel file 'AirQualityUCI.xlsx' using the Pandas library. The data contains information on various air quality parameters.

Data Preprocessing:

The dataset is loaded from the Excel file 'AirQualityUCI.xlsx' and split into training and testing sets. Feature scaling is applied to standardize the features to have a mean of zero and a standard deviation of one.

Gradient Descent:

Gradient descent is performed for each target column using the computed features and targets. The algorithm updates weights and bias iteratively to minimize the cost function, with optional L2 regularization (controlled by the lambda parameter).

Results:

The trained model is evaluated on the testing set, and the mean absolute error (MAE) is calculated for each target column. Additionally, the convergence of the cost function is visualized for each target column.

Conclusion:

In conclusion, this report demonstrates the use of Python programming for analyzing and visualizing air quality data.The gradient descent algorithm demonstrates its effectiveness in minimizing the cost function and generating predictive models for multivariate regression tasks. The reported MAE values provide insights into the model's accuracy and potential areas for improvement.

References:

Pandas Documentation: https://pandas.pydata.org/docs/

NumPy Documentation: https://numpy.org/doc/

Matplotlib Documentation: https://matplotlib.org/stable/contents.html