**Applied Machine Learning**

**Lab Report 6**

**Hafiz Ahmad**

**19l-1316**

**Section-8A**

**INTRODUCTION:**

Feature regularization plays a vital role in the analysis and modeling of data, especially when working with intricate datasets. Its purpose is to address the problems of overfitting and improve the overall performance of machine learning models by promoting generalization. The provided code exemplifies the application of feature regularization through the utilization of Ridge regression. Ridge regression is a commonly employed method that incorporates a regularization term into the linear regression model, allowing controlled reduction of feature complexity.

**OBJECTIVES:**

The main objective of this experiment is to apply feature regularization using Ridge regression to the 'Rented Bike Count' feature and analyze the impact of regularization on the feature.

**Procedure:**

Import Library: The code begins by importing specific scikit-learn modules and necessary libraries like numpy, pandas, and matplotlib.pyplot. Data manipulation, numerical computation, and visualization are all facilitated by these libraries.

Loading the Dataset: The read\_csv() function of Pandas is used to load the dataset from a CSV file. The segment names in the dataset are shown to give an outline of the accessible highlights. For the purpose of feature regularization, a list of the features of interest is created.

Information Planning: To set up the information for highlight regularization, the component values are removed from the dataset and put away in a numpy exhibit. Mean imputation is used to impute missing values from the array by replacing the missing values with the mean of each feature. This step guarantees that the information is finished and prepared for additional handling.

Scaling the Data: The StandardScaler from scikit-learn is used to scale the data. Scaling is important to make sure that all features are the same size so that no one feature can take over the regularization process. The StandardScaler changes the component values to have zero mean and unit difference, making them appropriate for regularization.

Regularization of Features: Ridge regression is used to apply feature regularization. With an alpha parameter of 1.0, the scaled feature array is used to fit the Ridge model. The "Rental Bike Count" feature is the focus of the regularization. The regularized values for this feature are then predicted using the fitted model.

Visualization: Using matplotlib.pyplot, the regularized "Rented Bike Count" feature can be seen. The index serves as the x-axis in a line plot, while the regularized feature values serve as the y-axis. This plot considers the assessment of the regularized example of the 'Leased Bicycle Count' highlight, giving bits of knowledge into possible enhancements with regards to diminishing intricacy and overfitting.

**Output:**

A picture containing text, font, screenshot, line

Description automatically generated

A picture containing text, font, screenshot

Description automatically generated

A picture containing text, font, screenshot, line

Description automatically generated

A black text on a white background

Description automatically generated with low confidence

A picture containing text, font, screenshot, line

Description automatically generated

A picture containing text, font, screenshot, line

Description automatically generated

A picture containing text, screenshot, plot

Description automatically generated

**Observations:**

The 'Rented Bike Count' feature was regularized by utilizing Ridge regression for feature regularization.

The regularized include values were acquired by fitting the Edge relapse model and anticipating the objective element.

Regularization may have reduced the original feature's noise or smoothed it out, as evidenced by the observed waveform- or noise-like patterns in the regularized feature values.

Overfitting can be reduced and the complexity of the feature reduced with feature regularization.

The regularization's nature and effect on the feature can be deduced from the waveform's specific characteristics.

**Application:**

Preventing overfitting: Regularization is a technique that manages the complexity of a model and enhances its ability to generalize by adding penalty terms such as L1 or L2 regularization.

Selecting relevant features: Regularization promotes sparsity, which helps in identifying important features and improves the interpretability of the model.

Reducing noise: Regularization methods, especially L2 regularization, reduce noise by shrinking the coefficients of less informative or noisy features.

Dealing with collinearity: Regularization addresses the problem of collinearity by distributing weights among correlated features, thereby improving the stability of the model.

**Issues:**

No issue was found while performing in the lab.

**Conclusion:**

Feature regularization using Ridge regression effectively reduces the complexity and overfitting in the 'Leased Bicycle Count' highlight. The waveform-like patterns in the regularized feature values indicate that the regularization has smoothed things out or reduced noise. A better feature's generalizability and robustness can be achieved through this regularization technique, increasing the accuracy and dependability of anomaly detection tasks.