**Applied Machine Learning**

**Lab Report 12**

**Hafiz Ahmad**

**19l-1316**

**Section-8A**

**INTRODUCTION:**

A technique called anomaly detection is used to find outliers or unusual patterns in a dataset. It can be used in a lot of different areas, like system monitoring, network intrusion detection, and fraud detection. We want to use the One-Class SVM algorithm in this experiment to find anomalies in the Seoul Bike Data. This dataset contains a few elements connected with bicycle rentals in Seoul.

**OBJECTIVES:**

The main goal of this experiment is to apply the One-Class SVM algorithm to the Seoul Bike Data and detect anomalies within the dataset. Additionally, we aim to analyze the characteristics of the identified anomalies.

**Procedure:**

Dataset and Library Import: We begin by loading the Seoul Bike Data and importing the necessary libraries.

Information Investigation: To see the attributes that are available, we print the dataset's columns.

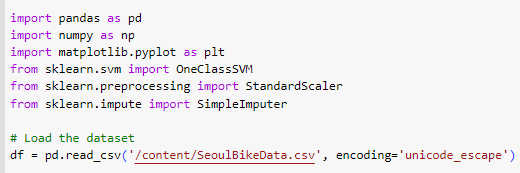
Information Planning: By altering the features to match the dataset's column names, the data is prepared for anomaly detection. Mean imputation is used to fill in missing values. We then apply include scaling utilizing the StandardScaler object. The fit\_transform method is used to scale and transform the data, resulting in a dataset that has been scaled and transformed.

SVM of One Class: We apply the One-Class SVM calculation for peculiarity location. An occurrence of the One-Class SVM classifier is made with a predefined boundary esteem (nu=0.05). The predicted labels for the data points are obtained after the classifier is trained on the scaled data.

Irregularity Perception: The data points are given labels for the anomalies, and scatter plots are made to show the anomalies. This step permits us to break down the peculiarities and their comparing highlights.

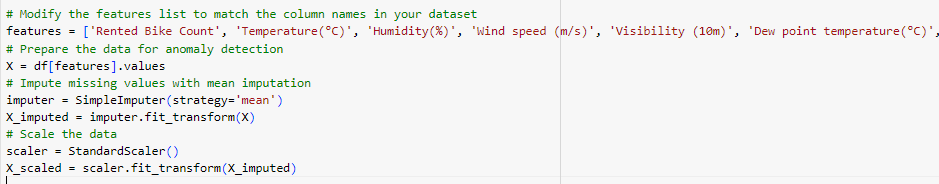
Spellbinding Insights: Expressive measurements are determined for the irregularity highlights. The descriptive statistics that were obtained by calling the describe() function on the "anomaly\_features" DataFrame are stored in the anomaly\_stats variable. The measurements, including count, mean, standard deviation, least, quartiles, and greatest, give experiences into the conveyance and qualities of the oddity highlights. In order to gain a deeper comprehension of the anomalies, box plots are also created.

Therefore, the libraries and dataset that will be used for anomaly detection will first be imported.



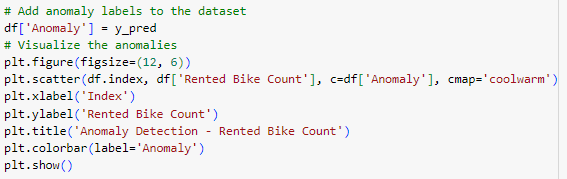
A picture containing text, screenshot, font, line

Description automatically generated



A picture containing text, font, screenshot, white

Description automatically generated



A screenshot of a computer code

Description automatically generated with low confidence A screen shot of a computer code

Description automatically generated with low confidence

A screenshot of a computer program

Description automatically generated with medium confidence

**OUTPUT By using the one class SVM:**

A picture containing text, screenshot, colorfulness, software

Description automatically generated

A screenshot of a computer

Description automatically generated with medium confidence A picture containing text, font, screenshot

Description automatically generated

A picture containing text, line, diagram, plot

Description automatically generated

**Results:**

We can observe the following from the output:

Results of Anomaly Detection: Out of the absolute 8760 significant pieces of information, 438 focuses are distinguished as inconsistencies, addressing around 5.00% of the dataset.

Anomaly Characteristics Descriptive Statistics: The table showcases graphic insights for each component of the identified inconsistencies. Remarkable perceptions incorporate the imply 'Leased Bicycle Count' of inconsistencies being roughly 592 with a standard deviation of 870, demonstrating sequential bicycle rental counts. Anomalies have a standard deviation of 13.66 and a mean temperature of 11.19, indicating that they occur at various temperature levels. Anomalies have a mean "Humidity(%)" of 64.40, with a standard deviation of 30.10, indicating that different humidity levels are associated with anomalies.

**Application:**

The One-Class SVM algorithm and other methods for anomaly detection are used in a variety of fields, including:

Detection of intrusions:

detecting potential security breaches or attacks by spotting unusual patterns in network traffic.

Detection of fraud: Identifying fake exchanges or exercises in monetary frameworks, for example, charge card extortion or protection misrepresentation.

Monitoring the equipment: identifying anomalies in the behavior of machinery or equipment to reduce the likelihood of malfunctions and improve maintenance schedules.

Wellbeing observing: Identifying unusual examples in clinical information, for example, important bodily functions or symptomatic test results, for illness discovery and patient checking.

**Issues:**

No issue was found while performing in the lab.

**Conclusion:**

We were successful in applying the One-Class SVM algorithm to the Seoul Bike Data for the purpose of anomaly detection in this experiment. Through the examination of the distinguished inconsistencies, we acquired significant bits of knowledge into their attributes. Anomaly detection results can be used to make informed decisions and identify unusual patterns with the help of these findings.