**Aim:**

To implement linear regression in statistical method

To implement linear regression in scikit learn

**Theory:**

**What is Linear Regression?**

**How do you compute the coefficients of regression line?**

**How do you check for the correctness of the model**

* **Karl pearson correlation coefficient**
* **R2**

**Solving a numerical on the same**

**Experiment:**

1. **Code in Python any data set with one dependent and one independent variable and Compute the following**
   1. **Coefficients and regression line**
   2. **The value r, SSE, R2**
2. **Code using sklearn using appropriate libraries linear regression algorithm using iris dataset and find the correlation between petal width and petal length and sepal length and sepal width for the 3 classes and consolidated for all the three classes.**

**Linear Regression Problem Definition:**

Predict the Petal Width, given: Petal Length

( Other simple linear regression and multilinear regression definition can be added after data analysis)

Compute the Data Analysis Report and tabulate the observations on Mean-squared error, R2 and Karl Pearson Coefficient for all the different combinations and analyse it in your own words.

|  |  |  |  |
| --- | --- | --- | --- |
| Combination taken | Dataset used | Metric used | Value |
| Petal Length and Petal width | Iris satosa | Mean-square error | 0.87(Eg) |
| Sepal Length and Sepal Width | Iris Satosa | R2 | (0.97)Eg. |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Analysis:

The correlation between Petal length and Widht seems so accurate since the sse is low & r and r2 are high.

**Conclusion:** The data analysis was performed on IRIS dataset and the following linear relationship was identified:

* Petal length and width are more correlated in comparison with the entire data set.
* Versicolor Petal length and width have high +ve correlation.