# MACHINE LEARNING PROJECT -08

SUBMITTED BY

DAVID SAMUELL J

195214108

I MSC COMPUTER SCIENCE

Problem Statement:

In this ML Project, I have used Iris dataset. Iris dataset has many features. In those features , I have chosen sepal length and sepal width feature and plotted in 2D graph using matplotlib. Then the Iris data is normalized using StandardScaler and values are predicted and their Mean Squared Error and Accuracy Score are predicted .

Machine Learning(Methodology):

The methodology used in this ML Project for training and testing the dataset is Linear Regression.

Linear Regression is a method to explain the relationship between a dependent variable and one or more explanatory variables using a straight line. It is a special case of regression analysis. This method is mostly used for forecasting and finding out cause and effect relationship between variables. Models depend linearly on their unknown parameters are easier to fit the models which are non-linearly related to their parameters.

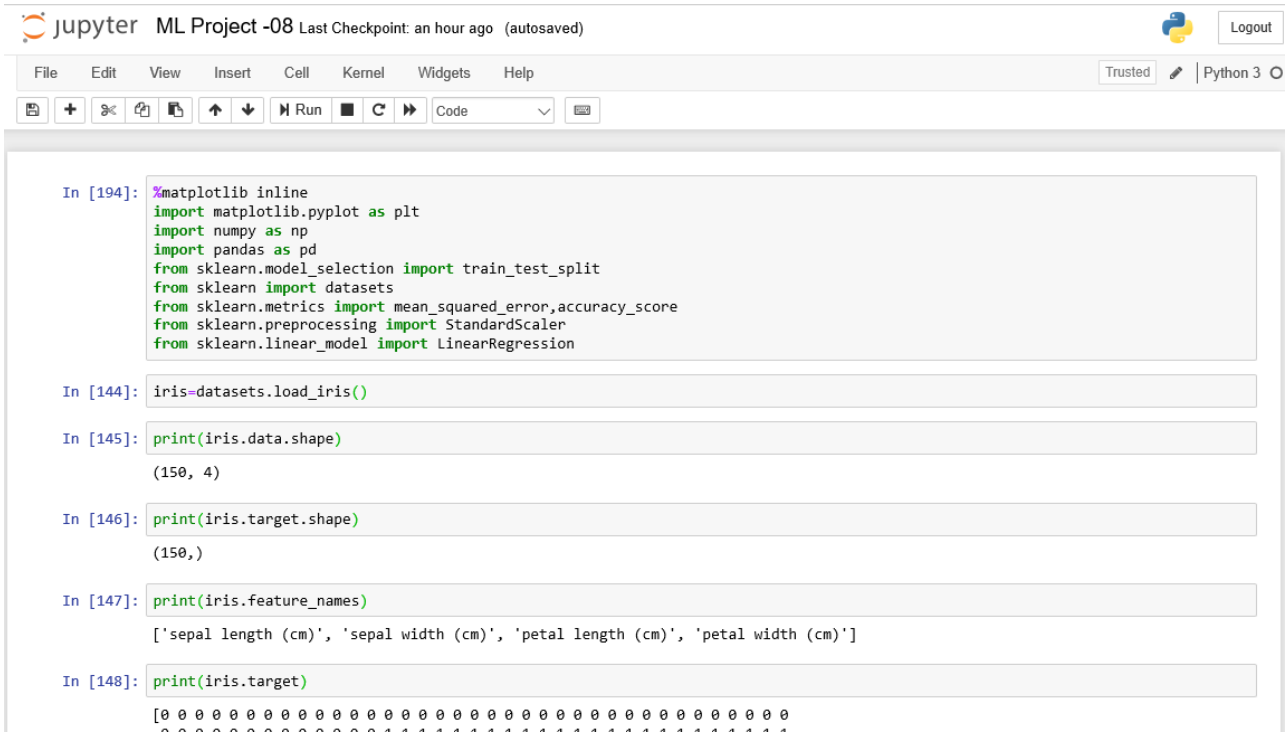
Dataset Description:

Some relevant columns in the dataset:

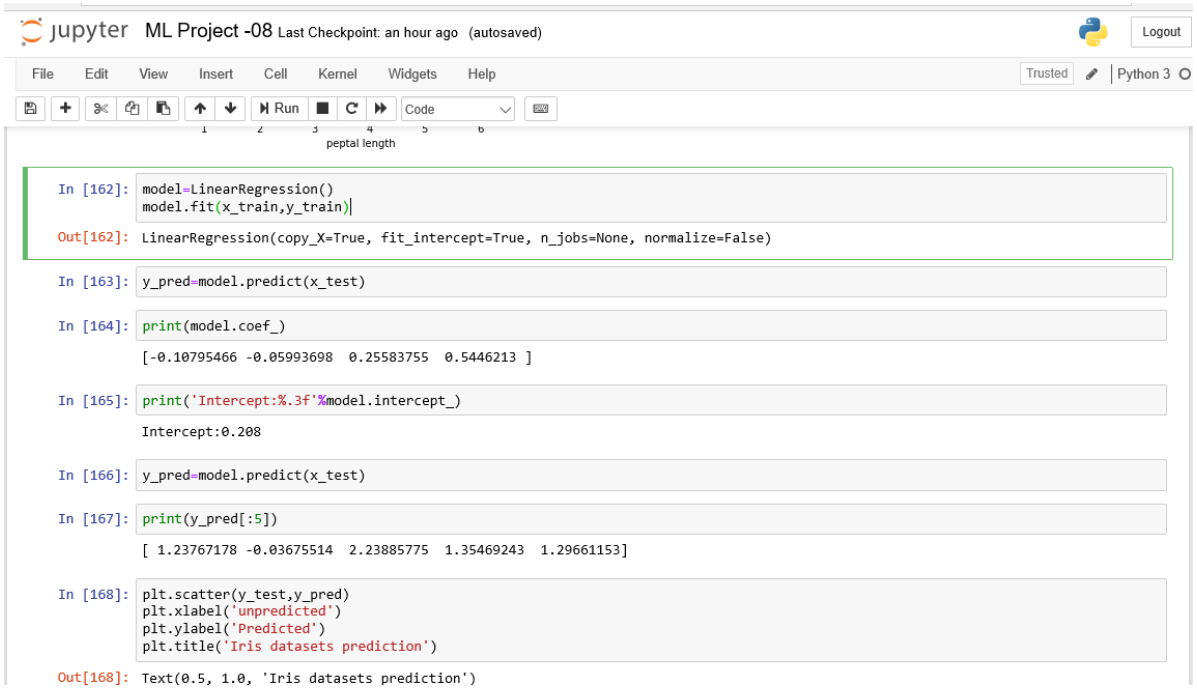
- Sepal Length
- Sepal Width
- Petal Length
- Petal Width

## Pre-Processing:

Pre-processing involves transforming raw data into an understandable format.  For example: Extracting data from a large dataset.




## Building ML Model:



# Training and Evaluation of all ML Models:

jupyter ML Project -08 Last Checkpoint: an hour ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

In [154]:
```
x_train,x_test,y_train,y_test=train_test_split(iris.data,iris.target,test_size=0.15,random_state=42)
```

In [155]:
```
scaler=StandardScaler()
x_train_std=scaler.fit_transform(x_train)
x_test_std=scaler.fit_transform(x_test)
```

In [156]:
```
print(x_train_std.shape)
```

```
(127, 4)
```

In [157]:
```
print(x_test_std.shape)
```

```
(23, 4)
```

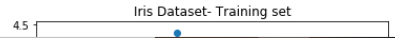
In [158]:
```
print(y_train.shape)
```

```
(127,)
```

In [159]:
```
print(y_test.shape)
```

```
(23,)
```

In [160]:
```
plt.scatter(x_train[:,:1],x_train[:,1:2])
plt.xlabel('Sepal length')
plt.ylabel('sepal width')
plt.title("Iris Dataset- Training set")
plt.show()
```




jupyter ML Project -08 Last Checkpoint: an hour ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

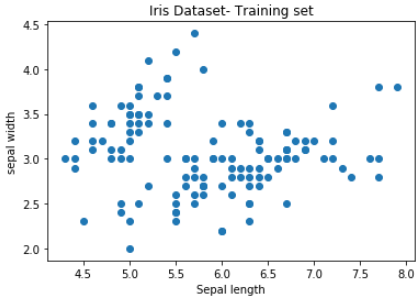
Run Code

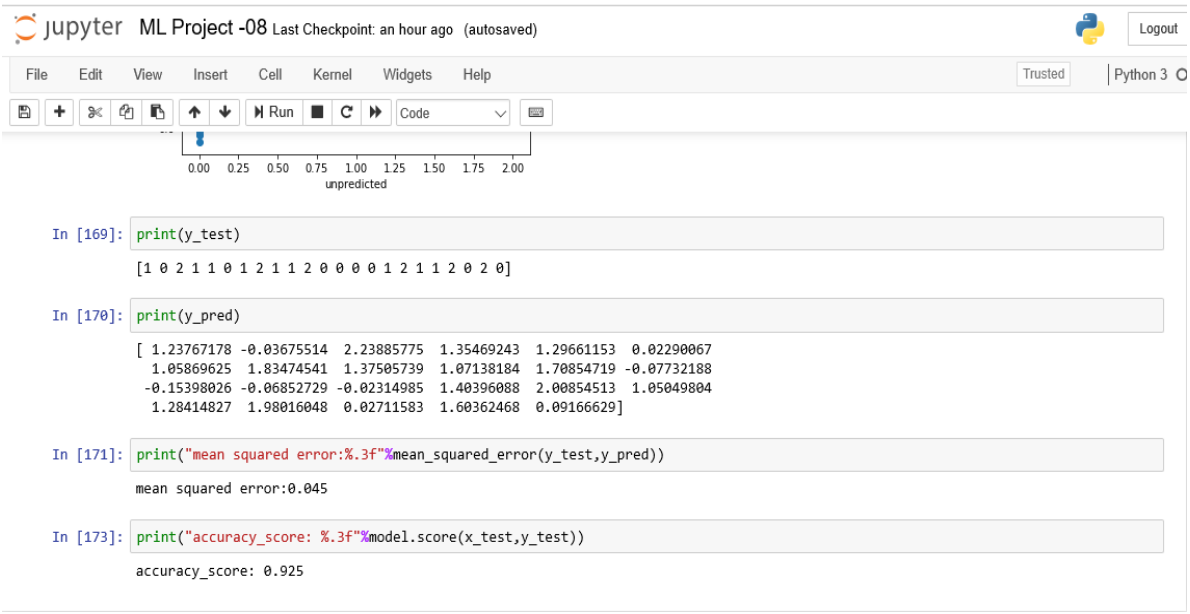
(127,)

In [159]:
```
print(y_test.shape)
```

```
(23,)
```

In [160]:
```
plt.scatter(x_train[:,:1],x_train[:,1:2])
plt.xlabel('Sepal length')
plt.ylabel('sepal width')
plt.title("Iris Dataset- Training set")
plt.show()
```






URL:

https://github.com/Davidsam98/ML-Project--08

Conclusion:

This Project gives the Predicted values of Iris Dataset Mean Squared Error and Accuracy Score using Linear Regression Algorithm.