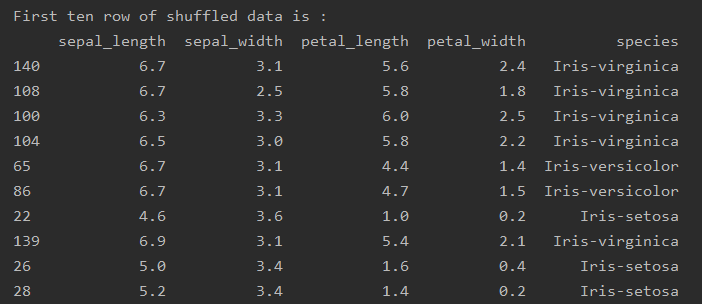
I used iris flower dataset for this project. In the dataset there are 150 samples with 5 features. Features are sepal length, sepal width, petal length, petal width and species of flowers. I used first 4 features to predict the last feature which is species of the flower. Since the dataset is in order i.e. first 50 sample belongs to iris setosa, second 50 sample belongs to iris versicolor and the last 50 sample of data belongs to iris virginica, I need to shuffle the dataset randomly.

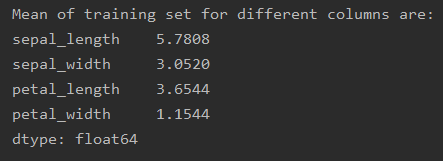
First ten row of shuffled data is:



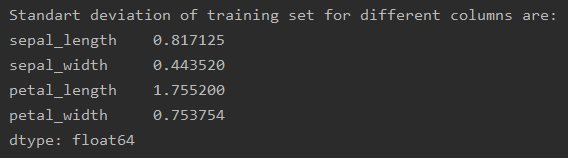
Species column will be used as labels of flowers, I will predict the species of given flower by using linear support vector machine classifier

Before training classifier, let get some information about data;

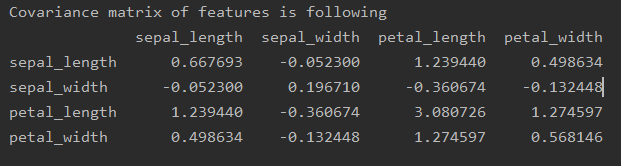
Mean of training set for different columns are;

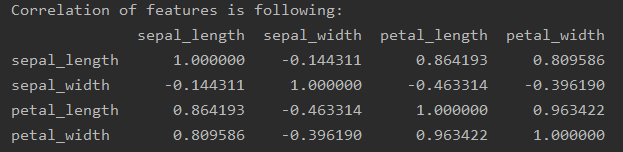


Standart deviation of training set for different columns are:



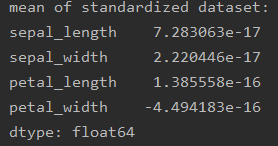
Covariance matrix of features is following

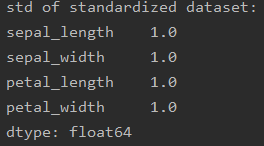
Correlation of features is following:



As it can be seen above correlation matrix there are high correlation between petal width and petal length as expected.

In order to standardize the dataset we need to get zero mean and one variance data so subtracting mean and dividing standard deviation gives standardized dataset;

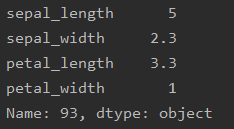




Means of standardized dataset are almost zero and stds of dataset are 1 as I want.

After fitting linear SVM classifier; we can predict outputs of given inputs

For example; Selecting a random data sample;



I order to put classifier I need to subtract mean which was calculated before and divide to std which was also calculated. Then putting the input to classifier gives output array;

[[-1.10178565 0.50027044 -2.01236246]]

Since maximum value(argmax) of the array is in 1. index and 1. element of the class is Iris-versicolor

It means the classifier predict the output of given flower as Iris-versicolor

Since the actual class of given flower is Iris-versicolor; prediction of classifier given randomly selected example is True.

After training the SVM the training accuracy over whole training set(125 samples) is; 94.39999999999999 %

Testing accuracy over whole test(25 samples) set is 96.0 %

It seems I get good test results, and the model predict species of given flower with high accuracy.