SUPPORT VECTOR MACHINES

Before starting SVM, you should accustomed with the Linear Regression and Logistic Regression.

SVM is first introduced in 1990’s and later got tuned in 1990.

SVM is a unsupervised Machine Learning Algorithm, which is used for both Classification and Regression. It will best for Classification problems.

SVM is classified into two types

* Linear SVM

In Linear SVM, where whole data point are perfectly separated into respective classes by using decision boundary

A hyperplane which maximize the margin between the classes is called decision boundary.

A straight-line or hyper-plane is enough to divide the datapoints into their respective classes.

* Non-Linear SVM

It is used to classify the data, where the datapoints cannot be separated to their respective classes by using a straight-line or hyper-plane.

In these scenario kernel function comes into picture. It transforms the original input data into high-dimensional feature space, where the data points are linearly separated.

The main ideology of SVM is to find the hyperplane in a N-dimensional space that can separate the data points into their respective classes.

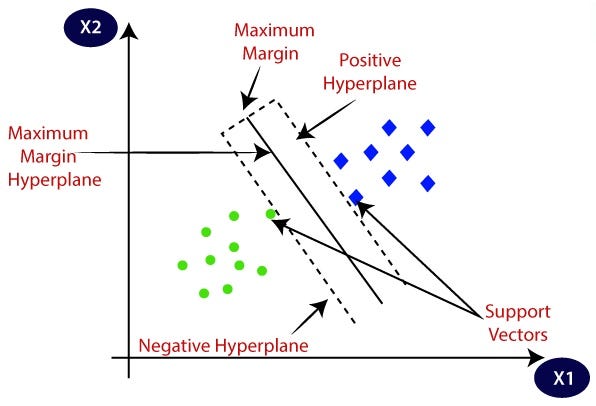
If the Number of input features are two, then the hyperplane is 1-D (Straight-Line)

If the Number of input features are three, then the hyperplane is 2-D.

If the Number of input features are four, then the hyperplane is 3-D.

**SVM Working Procedure**

Let us consider two scenarios to know the working procedure of Support- Vector Machine



From the above figure, we can see that there are two classes i.e green class and blue class. These both the classes are separated by a line. There are 3 different lines in the figure, where they separate both the classes, but the best fit line is the middle most line among three. Because that middle most line is called the “**maximum-margin hyperplane/hard margin”, since the nearest datapoint on each side is maximized.**

**When the line separate positive class from the negative classes, then it is said to be Positive Hyperplane. Mathematically, it is denoted by +1.**

**When the line separate negative class from the positive class, then it is said to be Negative Hyperplane. Mathematically, it is denoted by -1.**



Here, there are two classes one is red class and other is blue class. So in the above figures a blue data point is mixed with red class. How SVM will classify the data? It’s Simple! The blue ball in red ball boundary is an outlier. SVM has a characteristic to ignore the outliers and find best hyperparameter that maximizes the margin.

SVM is Robust to Outliers.

Till now we have discussed about linear SV. Now we will discuss about Non-linear SVM. Below is the example for non-linear SVM, where the data points are misclassified. How SVM solve this type of data? It’s Simple!!



Consider a input as xi and other input yi which is shown in the figure and plot the data points according to the original data in the feature space of xi and yi the try to implement a hyperplane which maximize the distance from both the data points.



**Terminology of SVM :**

**Hyperplane:** A Straight line which divides the data points into their respective classes in a feature space.

**Support Vectors:** Support vectors are the data points which are near to the hyperplane.

**Kernel:** Kernel is a mathematical function, which is used for the data which are misclassified. The main function of kernel is, it transforms the original data points in to a high-dimensional feature space, So that the hyperplane can be easily found out even the data points are misclassified. Some of the kernel functions are

* Linear
* Polynomial
* Gaussian Rbf
* Sigmoid

**Soft Margin**: It is a technique to solve the misclassification or Non-Linear data. When the data is not separated correctly SVM allows Soft-Margin technique. It softens the stract margin requirements and permits certain misclassification and margin violations.

**Hard Margin:** It is also a hyperplane which does not allow miss-classifications. It separates the data points perfectly, by having a best fit line in between the data points.

**Hinge Loss:** It is a typical loss function. It punishes incorrect misclassifications and margin violations.

**C**: Margin maximization, misclassification fines are balanced by the regularization parameters C in SVM. Penalty for going over the margin or misclassified data items is decided by it. Greater the value of C, higher the penalty.

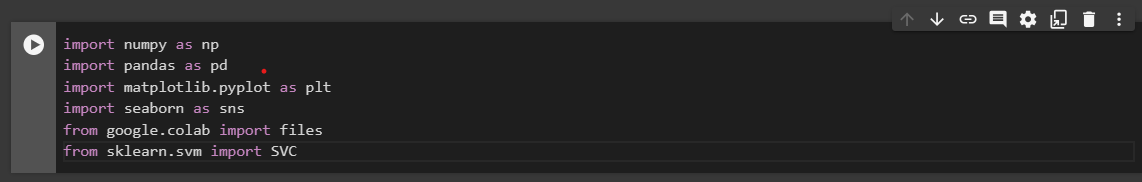
**Practical Implementation:**

Predict whether the given data is Setosa, Virignca or VersiColor

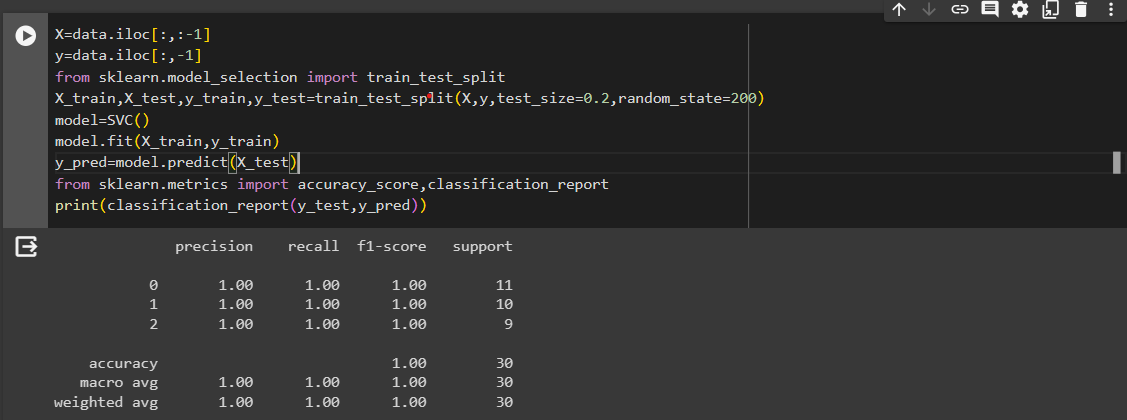
Steps to follow:

* Import the necessary Libraries to implement SVM Algorithm.
* Upload the data from local computer or from any other source and read the dataset.
* Apply data cleaning and identify the patterns in the data.
* Separate input and target variable.
* Build and train the SVM Classifier.
* Plot the scatter plot of input features.
* Test the SVM Classifier model.
* Validate the Model.

First to implement the Support Vector Machine Algorithms, we need to import necessary libraries.



The given data is Iris data and imported necessary libraries and the next step is to upload the data and check whether the data is cleaned and having any trend or pattern and then forward to feature engineering and after that split the data into training and testing and apply Support Vector Machine Algorithm and check the accuracy of that model given to the respective dataset.



Here the model is giving the accuracy as 1 that is 100% it is predicting the correct result.

**THANK YOU**