**AISC1003 - Machine Learning 1**

**2021F**

**Application Exercise 3**

**Report on Support vector machine**

**Instructor: Prof. Jagmohan Dutta**

**Name : Pratik Domadiya (ID : 500199494)**

1. **Introduction**

SVM : support vector machine

SVM is a popular supervised algorithm to use in the task of regression as well as classification problem in machine learning, though it is widely used for classification.

The main objective of the SVM is to find out hyperplan that classifies the data.



As seen in the illustration, there are two types of data available: one with a circle and the other with a square. We can now see that there are an endless number of hyperplanes (here 1d line) that can distinguish two classes. As a result, the SVM's purpose is to discover the best hyperplan for dividing a given collection of data.

Source : [https://miro.medium.com/max/600/0\*9jEWNXTAao7phK-5.png](https://miro.medium.com/max/600/0*9jEWNXTAao7phK-5.png)



As we can see an optimal hyperplan which can divide the two different sets of the data. Here a term “**maximum margin**” is defined as a maximum distance between the data points of the both classes.

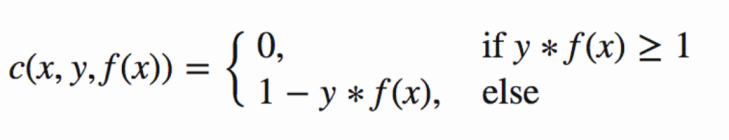
Source : [https://miro.medium.com/max/600/0\*0o8xIA4k3gXUDCFU.png](https://miro.medium.com/max/600/0*0o8xIA4k3gXUDCFU.png)

Support vector machine widely used in the complex machine learning task such as handwriting recognition, face detection, email classification and gene classification.

If you have data with n different target classes then SVM will give you n-1 dimensional hyperplan in output. So end goal of the SVM is to maximize the margin of the classifier in order to get well and accurate prediction.

**COST FUNCTION** :

Goal : maximize the margin



Source : [https://miro.medium.com/max/2080/1\*3xErahGeTFnbDiRuNXjAuA.png](https://miro.medium.com/max/2080/1*3xErahGeTFnbDiRuNXjAuA.png)

We can also add the regularization parameter into the cost function. The objective of the adding regularize term is to balance the margin maximization and loss.

1. **Different types of the kernel in SVM**

Kernel plays a crucial role in classification to identify the hidden pattern from the data. It

Also helps with high dimensional data in a very efficient manner.

* Linear Kernel

: Linear kernel is the most basic type of kernel. It usually one dimensional in the nature . When there are lots of features are available in the data, linear kernel works the best.

: **F (x,xj) = , where** x,xj is the data you are trying to classify.

* Polynomial Kernel

: Polynomial kernel is the more generalized representation of linear kernel. But it is used less as compare to other types because of it is less efficient and accurate.

: **F (x,xj) =, where** x,xj is the decision boundary to separate the given classes and d defines the degree.

* RBF – Gaussian Radial Basis Function

: RBF is the most preferred and used kernel functions in SVM. It is very helpful when we have a non-linear data.

: **F (x,xj) =, where** 0 < gamma <1, the most preferred value for gamma is 0.1.

* Sigmoid Kernel

: When a case of neural network, sigmoid works better than other types.

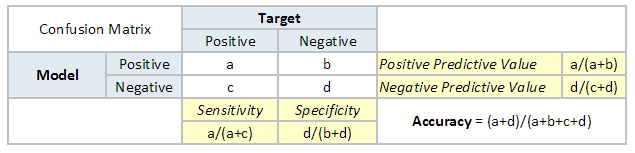
* Gaussian Kernel

: Gaussian kernel used when there is no prior knowledge of the given dataset.

: **K(x,y) =**

1. **Metrics For SVM**

* Confusion metrics : It is used to find amount of values which are predicted correctly and wrongly.

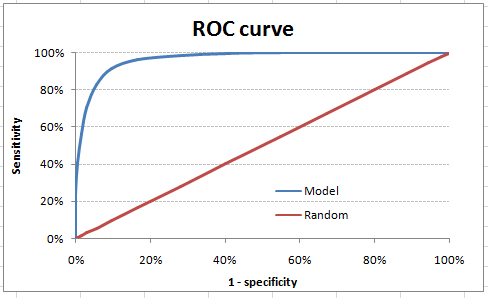


Source : <https://www.analyticsvidhya.com/wp-content/uploads/2015/01/Confusion_matrix.png>

* False positive is considered as a type 1 error
* False negative is considered as a type 2 error
* Classification Report:
* Precision : TP / (TP + FP)
* Recall : TP / (TP + FN)
* F1 score : Harmonic mean between precision and recall

: 2\*TP / (2\*TP + FP + FN)

* Accuracy : (TP + TN) / (TP + FN + FP + TN)
* Predict the overall accuracy of the model.
* ROC curve :
* **R**eceiver **O**perating **C**haracteristics curve. The roc score lies between 0 and 1. If the score is 0.78 then it means 78% of the predicted values are correct and rest 22% are wrongly predicted.



Source : <https://www.analyticsvidhya.com/wp-content/uploads/2015/01/ROC.png>

1. Assumption of SVM

There are mainly two assumption available for SVM :

* A few data values are independent
* A minimum data values are identically distributed.

References :

<https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>

<https://dataaspirant.com/svm-kernels/>

<https://www.analyticsvidhya.com/blog/2019/08/11-important-model-evaluation-error-metrics/>

<https://towardsdatascience.com/20-popular-machine-learning-metrics-part-1-classification-regression-evaluation-metrics-1ca3e282a2ce>