# Support Vector Machine

* Supervised multi-class classifier model
* Can be used for both classification and regression. Mostly used in classification problems
* In this algorithm we plot each data item in n-dimensional space (n is number of features) with values of each feature being the value of particular coordinate.
* Then, we perform classification by finding a hyper-plane that differentiates the two classes very well.
* Margin: maximize the distance b/w nearest data point & hyper plane. Margin will help us decide the right hyper plane.
* SVM ignores the outliers. This is a main advantage
* Selects the hyper-planes which classifies the classes prior to maximizing margin.
* **Kernel** Trick: Here the data is take from lower dimensions to higher dimensions, so its each to chop or classify the data.
* **Model Parameters:**
  + **Gamma:** higher the value of gamma, will try to exact fit as per training data. May causes overfitting problem.
  + **C-penalty parameter**: controls the trade-off b/w smooth decision boundary and classifying the training points correctly.

**Model performance:**

* We will perform cross-validation, split the data into train and test datasets (70-30). Build a model on training dataset and check the training accuracy.
* Now apply the model on test dataset (using predict function). Check test accuracy.
* If the train and test accuracy is not satisfactory, then change the model parameters and build new models.
* Continue this step until u get a decent train and test accuracies.(By doing this you will not over fit the model on train dataset)
* Also check precision, recall, F1 score and Support.
* Finally we can check explained variance value (best is 1) and mean absolute error (best is 0)

**NOTE: This should be performed for all the models. K- Fold validation can also be performed.**