**What is SVM:** Supervised learning method which can be used for both classification and regression. However, it’s mostly used in classification problems.

**How is the data classified:** We perform classification by finding the hyperplane that differentiates the two classes very well. In other words, the algorithm outputs an optimal hyperplane which categorizes new examples.

**What is an optimal Hyper-Plane:** For SVM, it’s the one that maximizes the margins from both tags. In other words, the hyperplane whose distance to the nearest element of each tag is the largest.

A diagram of a graph

Description automatically generated with medium confidence

**Nonlinear data**

**Hyperparameter Tuning:**

Kernel: The learning of the hyperplane in linear SVM is done by transforming the problem using some linear algebra. This is where the kernel plays role. Polynomial and exponential kernels calculate separation line in higher dimension. This is called kernel trick.

Gamma: The gamma parameter defines how far the influence of a single training set reaches. With low gamma, points far away from the possible separation line are considered in calculation for the separation line. Whereas high gamma means the points close to possible line are considered in calculation.

Regularization: For large values of this parameter, the optimization will choose a smaller-margin hyperplane if that hyperplane does a better job of getting all the training points classified correctly. Conversely, a very small value of it will cause the optimizer to look for a larger margin separating hyperplane, even if that hyperplane misclassifies more points.

Margin: A margin is a separation of line to the closest class points. A good margin is one where this separation is larger for both the classes. A good margin allows the points to be in their respective classes without crossing to other class.