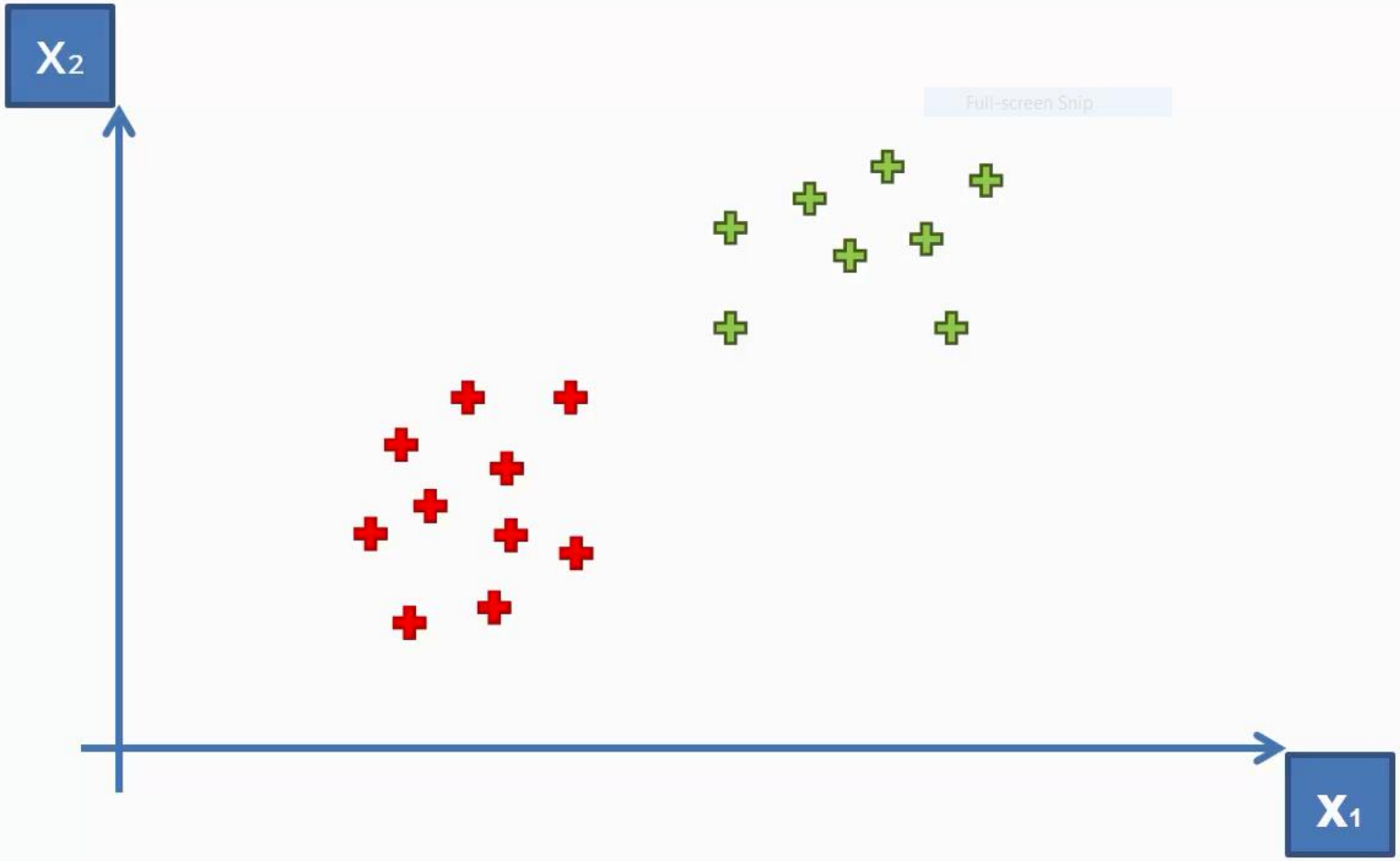
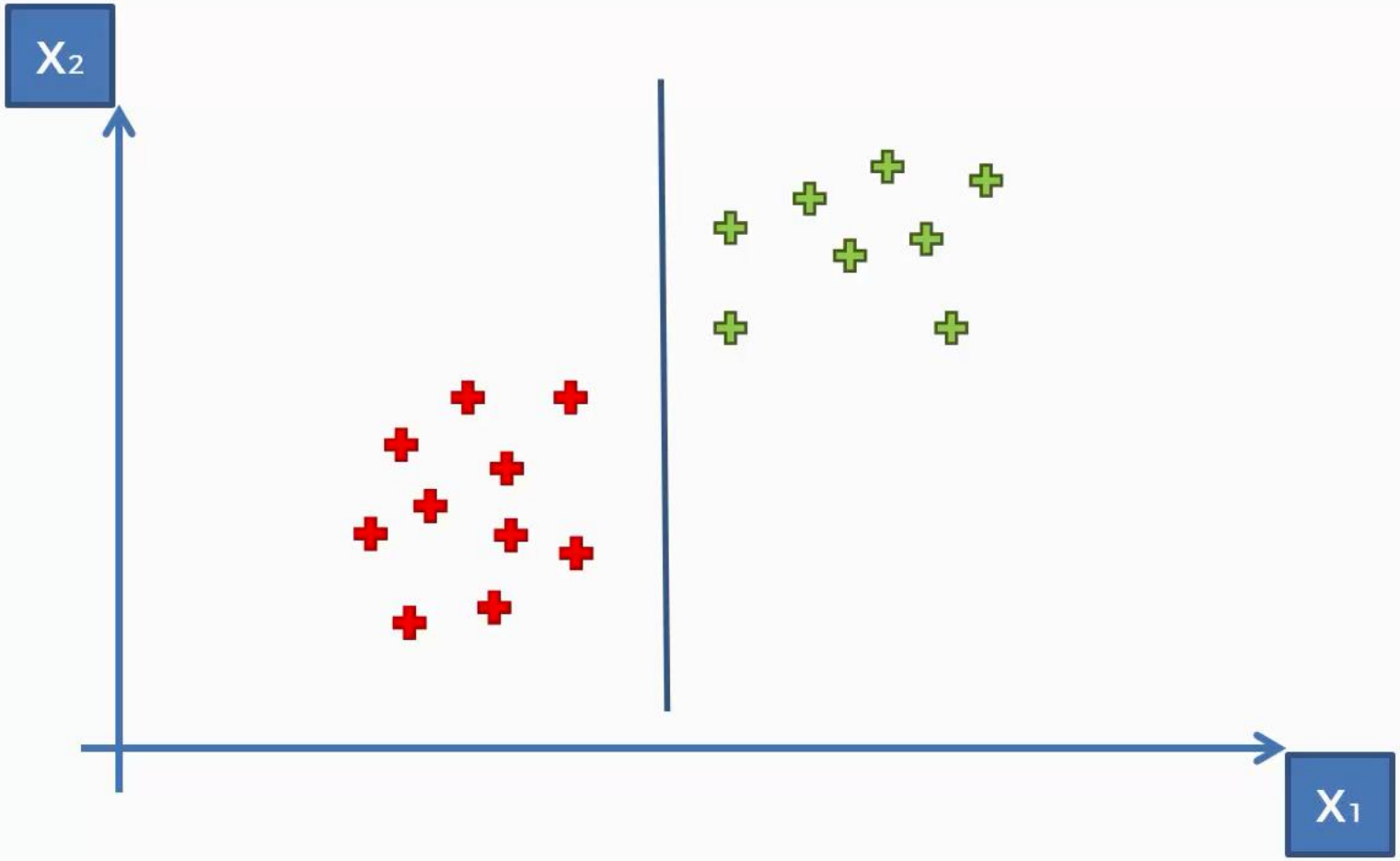
Support Vector Machines (SVM)

Non Parametric, Binary Classification Technique

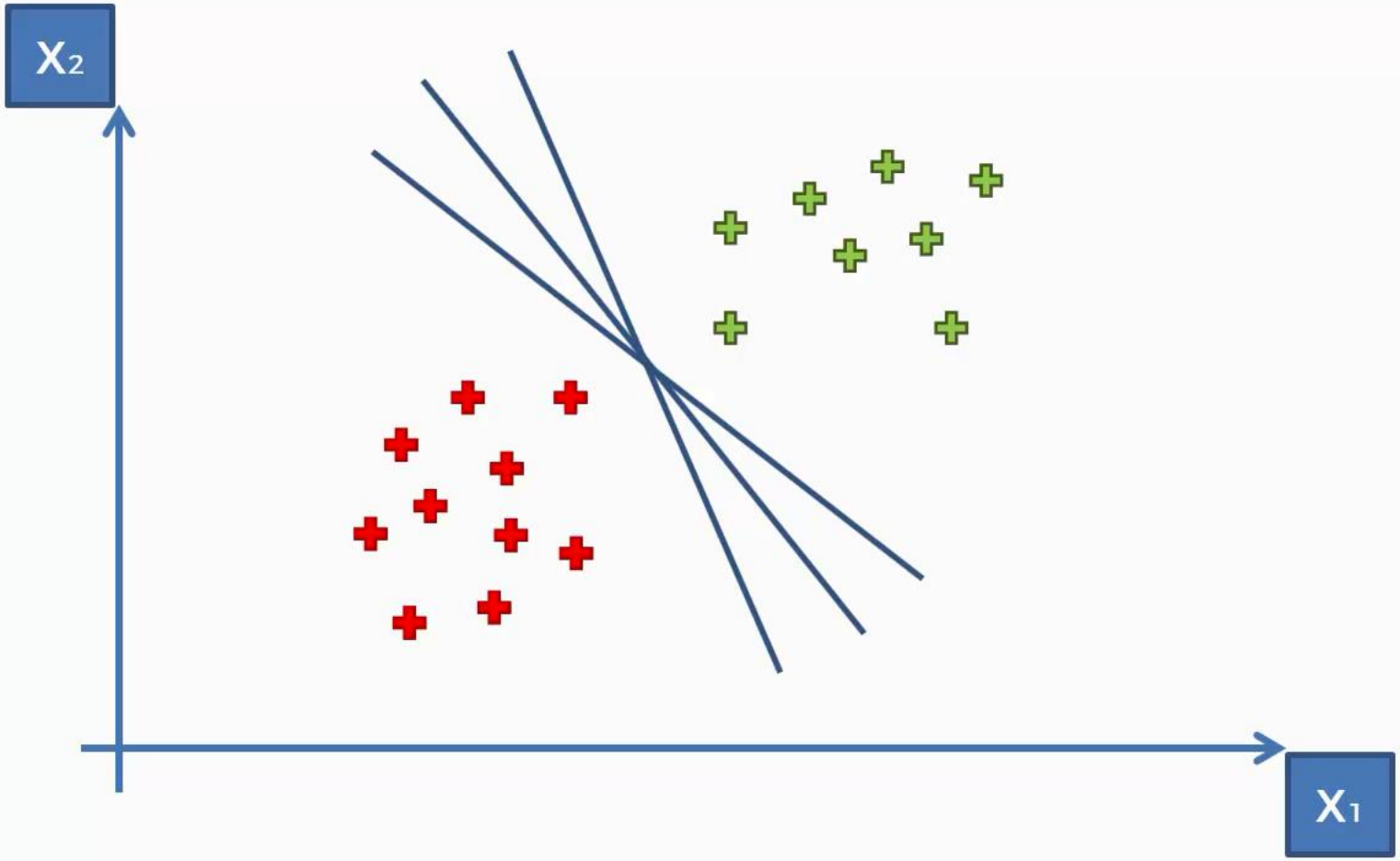
How to separate these points ?



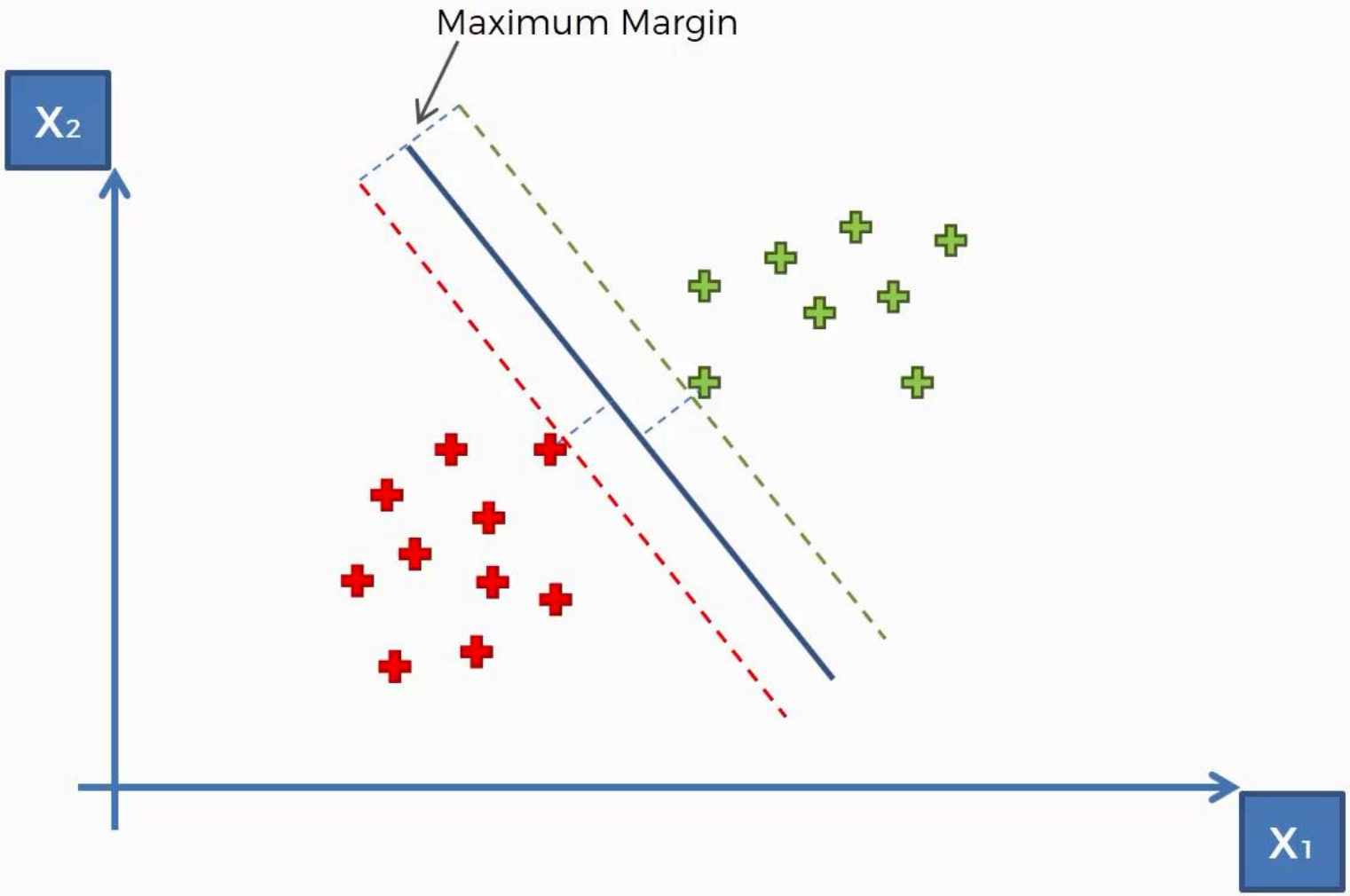
How to separate these points ?



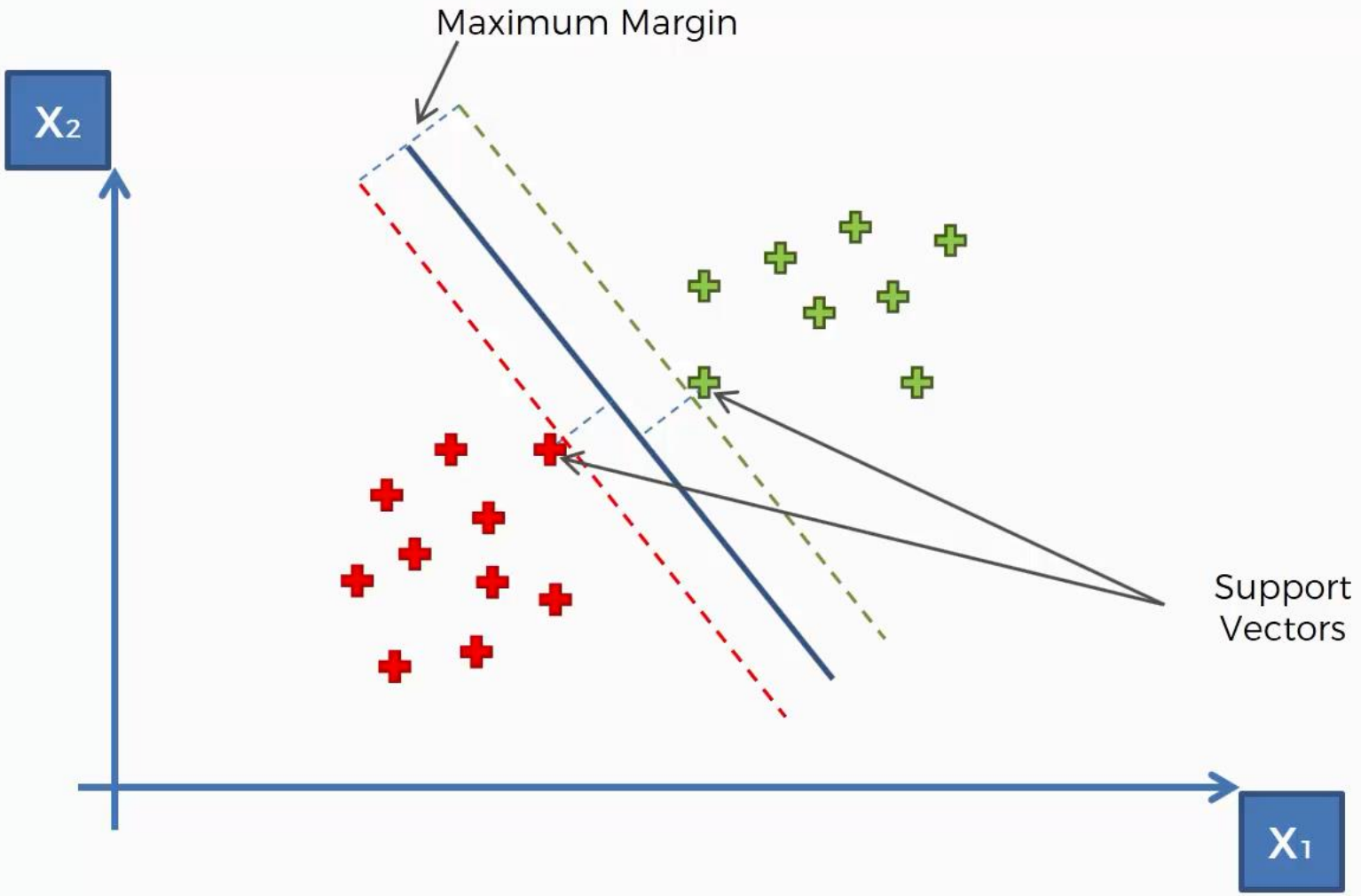
SVM



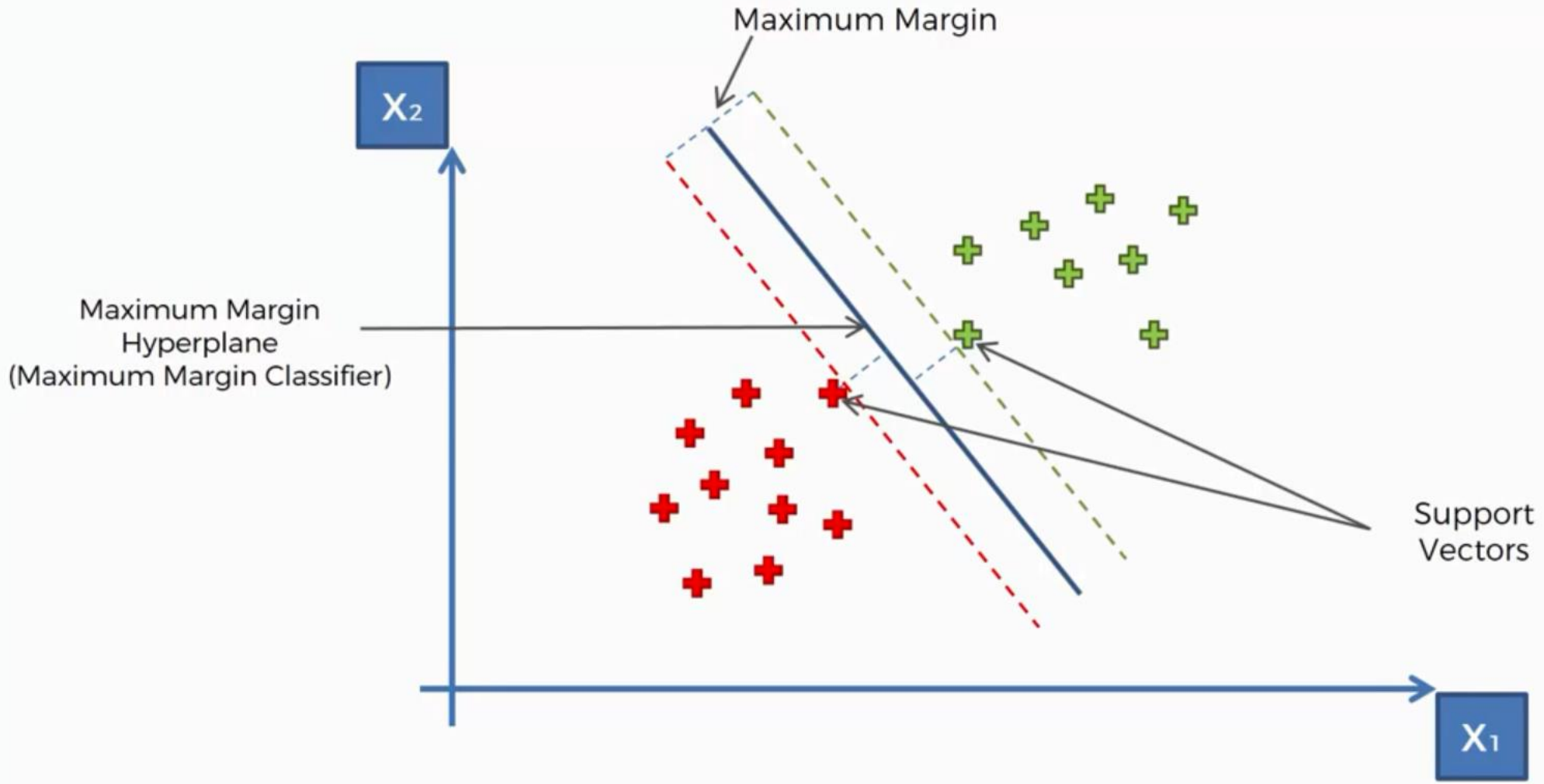
Maximum Margin



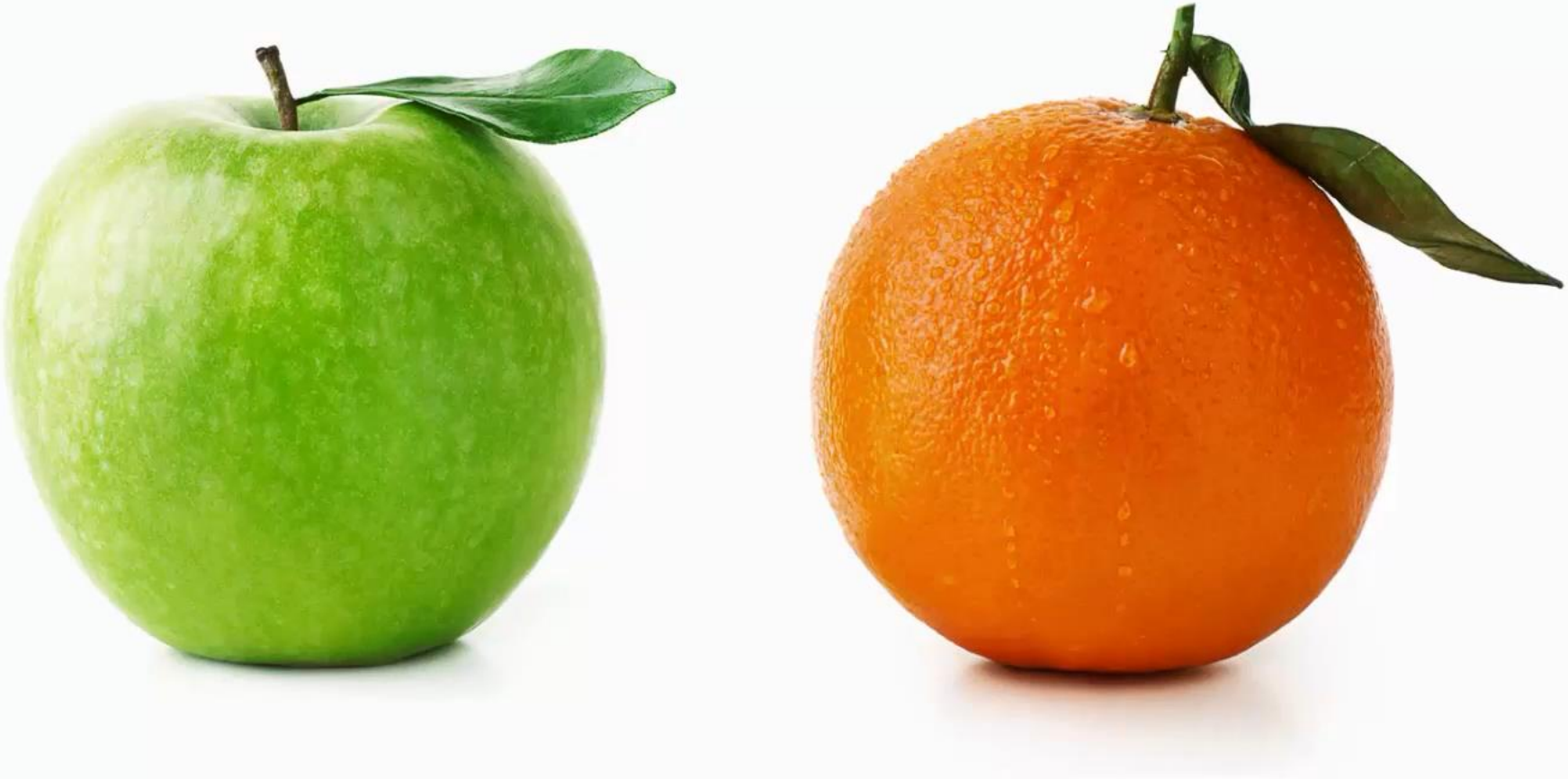
Support Vectors



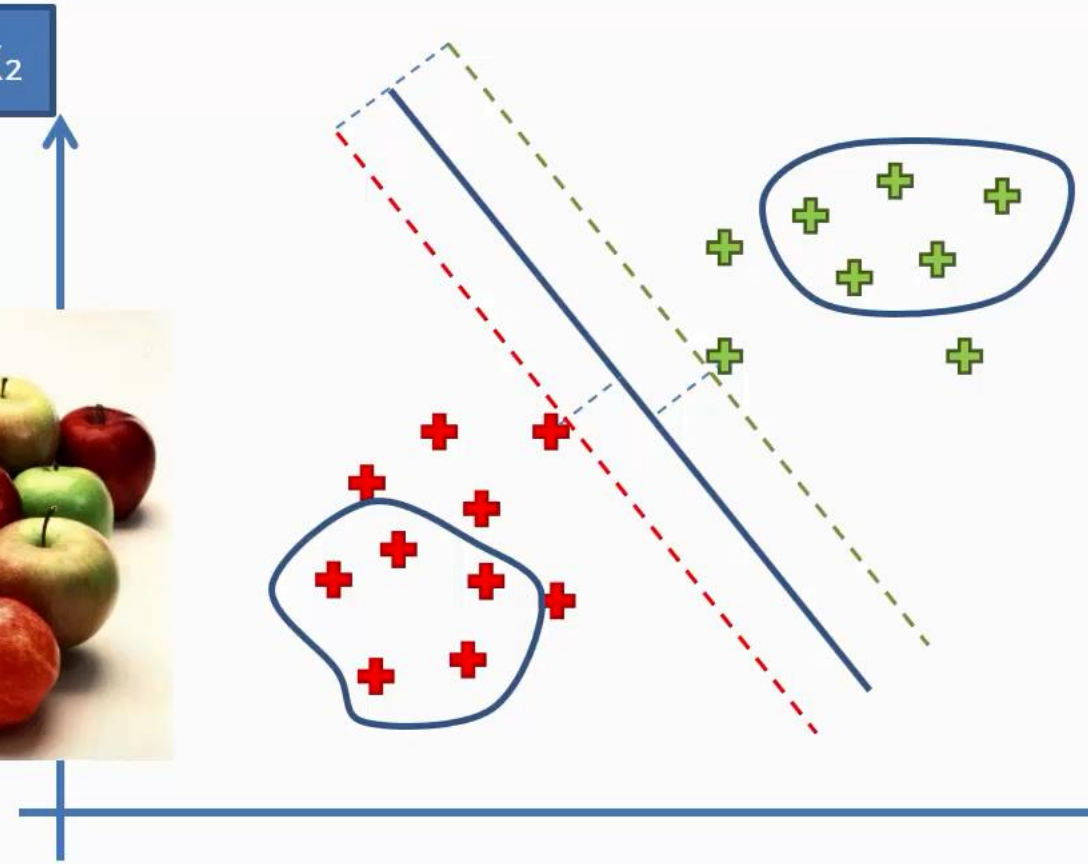
Hyperplanes



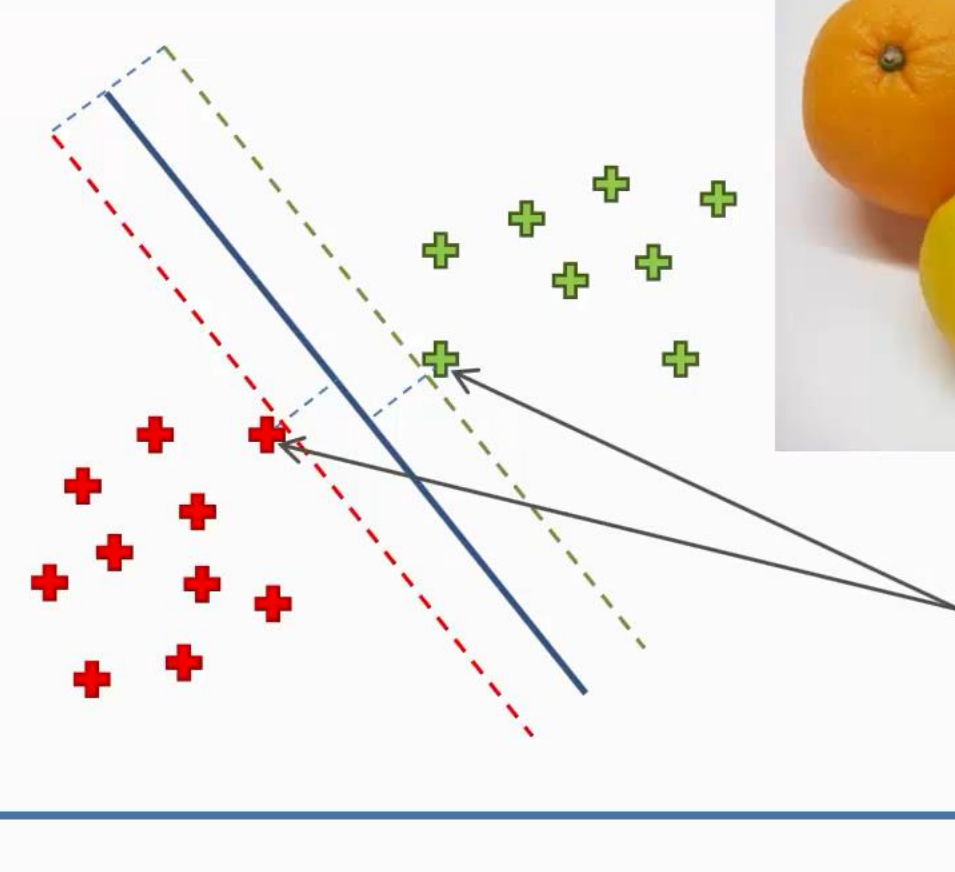
What's So Special About SVMs?

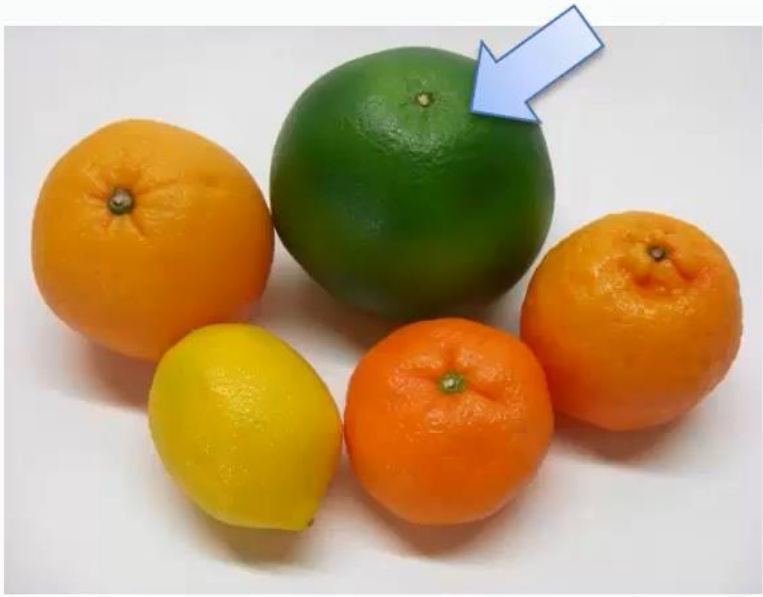


What's So Special About SVMs?



What's So Special About SVMs?





Model Tuning

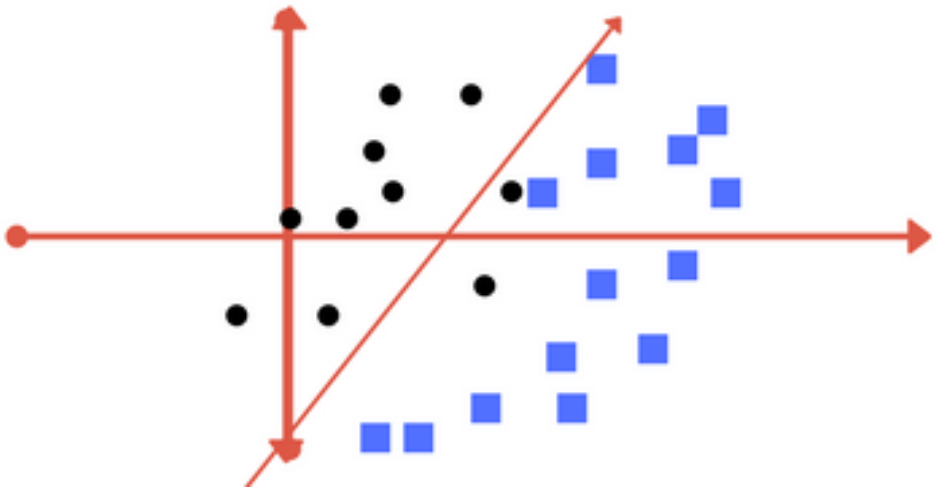
Tuning SVM Parameters to obtain better results.

Regularization (C Parameter)

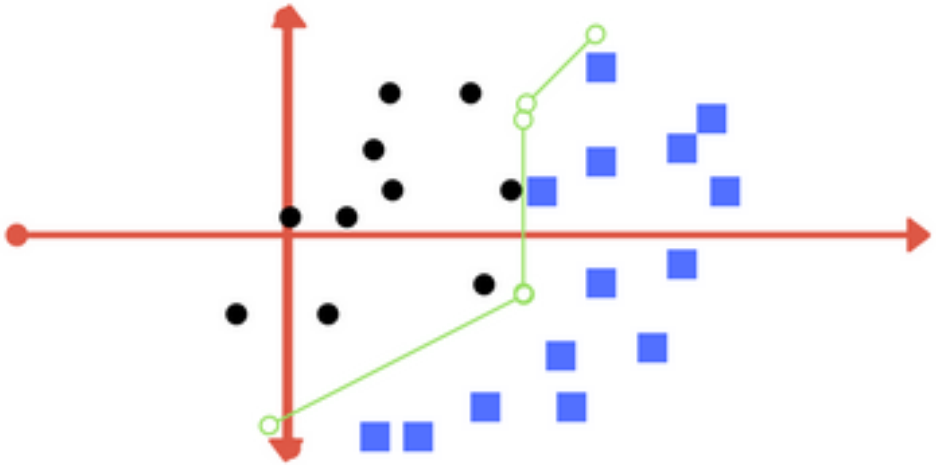
The Regularization parameter (often termed as C parameter in python's sklearn library) tells the SVM optimization how much you want to avoid misclassifying each training example.

For large values of C, 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 C will cause the optimizer to look for a larger-margin separating hyperplane, even if that hyperplane misclassifies more points.



Low Regularization (C) value



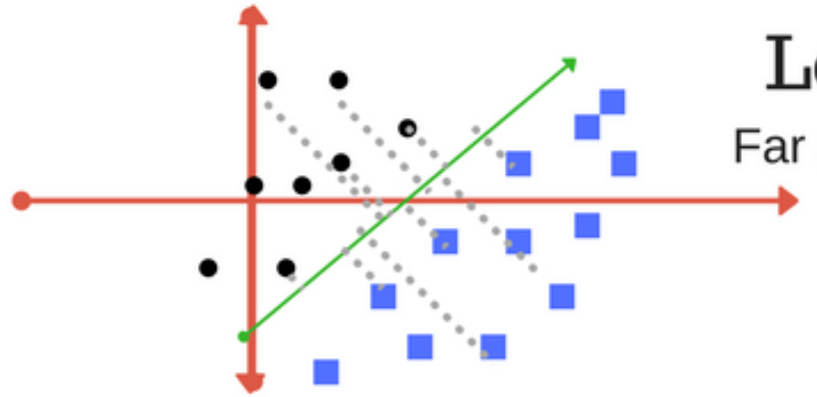
Higher Regularization (C) value

Gamma

The gamma parameter defines how far the influence of a single training example reaches, with low values meaning ‘far’ and high values meaning ‘close’.

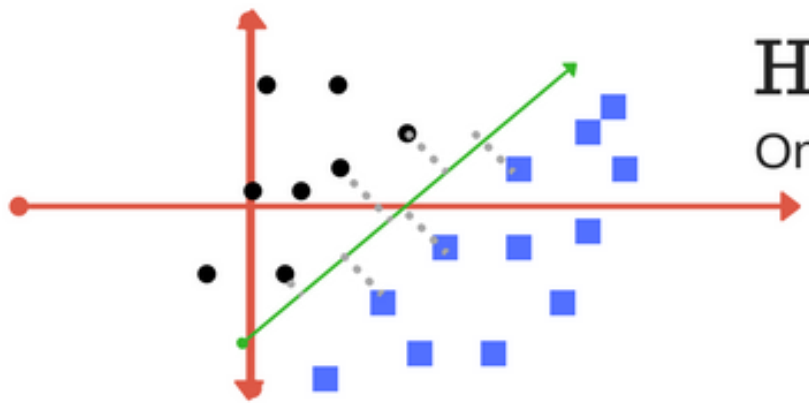
For lower gamma value, the points far away from plausible separation line are considered in calculation for the separation line.

Higher gamma means the points close to plausible line are considered in calculation.



Low Gamma

Far away points are also considered.



High Gamma

Only nearby points are considered.