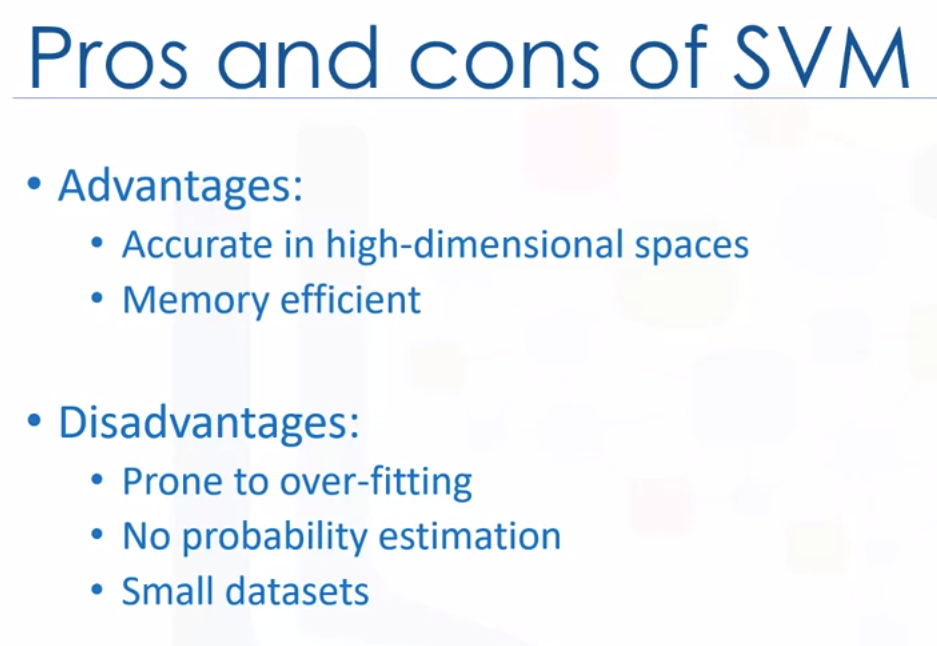


Maximize margin

Kernelling in SVM: Mapping data into a higher dimensional space, in such a way that can change a linearly inseparable dataset into a linearly separable dataset.



Adv: they use a subset of training points in the decision function called, support vectors, so it’s also memory efficient.

Dis: The algorithm is prone for over-fitting if the number of features is much greater than the number of samples.

SVMs don’t directly provide provide probability estimates, which are desirable in most classification problems. And finally, SVMs aren’t very efficient computationally if your dataset is very big, such as when you have more than 1000 rows.

