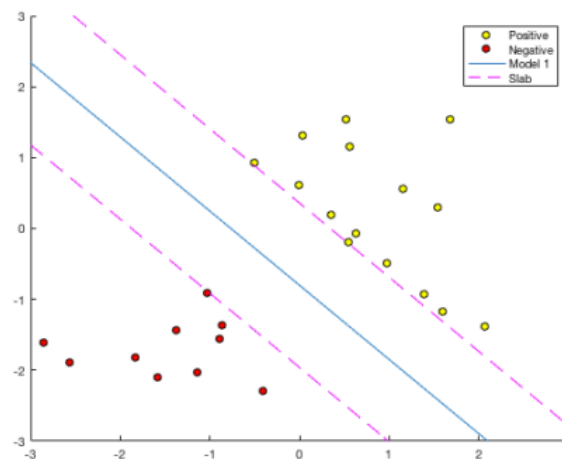


Maximal margin classifier and SVM

Hyperplanes: are a geometric structure, in particular a subspace of dimension one less than its ambient space.

we want to find a hyperplane that separates our data and choose which has the largest margin separating the two classes.



The support vectors are the data points close to the hyperplane/decision boundary that contributes to its best position.

How do we decide how to transform the data? To make mathematics possible, SVMs use something called **Kernel Functions** to *systematically* find SVC in higher dimensions.

What happens when there is no clear separating hyperplane (kernel SVM)?

If we have a **linearly separable dataset** then the SVMs job is usually easy. However, in real life, in most cases, we have a linearly non-separable dataset at hand and this is when the **kernel trick** provides some magic.

The **kernel trick** projects the original data points in a higher dimensional space to make them linearly separable (in that higher dimensional space).