Support vector machine (SVM) models are classification models, mainly used for two labels, that assume linear separability in order to predict labels. The general idea is to create an object that lives one dimension lower than that of the data. So, for a typical scatterplot with two dimensions, we would draw a line to separate the classes. As dimensionality increases, the separation object increases in complexity as well.

A possible problem with this method is that data may not be linearly separable. To counteract this issue, we can use several tricks or methods to separate the data. One important point to raise here is the use of soft and hard margins. A hard margin assumes linear separability, while a soft margin assumes it’s not possible and allows data points to cross the hyperplane. Soft margins then assume that the great majority of points can be classified as initially intended, but with a few errors.

Furthermore, the Kernel Trick can also be used on on-linearly separable data. The Kernel Trick maps the data inputs onto a higher dimension to create an object that could successfully separate data that on a lower dimension is not separable.