

A support vector machine is a supervised machine learning algorithm that can be used for classification or regression purposes.

Because it is commonly used for classification, that will be our focus.

Generally speaking, SVMs are used to find a hyperplane that best divides a multi-dimensional dataset into two classes.

Support vectors are the data points nearest to the hyperplane that separates the two classes.

These data points are considered critical elements because if they were removed, it would alter the position of the dividing hyperplane.

In a simplified case of only two features or a two dimensional space, the hyperplane becomes a line that separates a dataset into two categories or classifications.

From this point on, we will use this simple case.

The further the points are from the hyper plane, the more certainty we can have that the data points from the training data are classified correctly.

Using this information, we want our test data points as far from the hyperplane as possible, but still on the correct side.

Then when we add new data, we have greater confidence that the new data points are classified correctly.

The distance from the hyperplane and the closest data point is called the margin.

We want to choose a hyperplane that has the largest margin between itself and any data point in the training data.

This will give the best chance that any new data will be classified correctly.

However, data is usually tangled together or rarely clean.

This means that it is often hard to place on a line. To better see the classification, it's necessary to display this in a 3D view.

Imagine that are two sets of colored data points are sitting on a sheet that is lifted, throwing the balls into the air. While the balls are up in the air, you use the sheet to separate them.

This lifting of the balls represents the mapping of data into a higher dimension, and this is known as kerneling.

Because we are now in three dimensions a hyperplane can no longer be aligned.

It must now be a plane as shown.

The idea is that the data will continue to be mapped into higher and higher dimensions until a hyperplane can be formed and segregated.

SVM is used for text classification tasks, such as category assignment, detecting spam, sentiment analysis.

It is also commonly used for image recognition challenges, performing particularly well in aspect based recognition and color based classification.

SVM also plays a vital role in many areas of handwritten digital recognition, such as postal automation services.

So there you have it, a very high level introduction to support vector machines.

We will, of course, be using these concepts throughout the remainder of the course and always thank you for watching.