

### Support vector machines

Classifier learning can be seen as finding a curve that well separates two sets of points. We are trying to find an optimal curve out of infinitely many curves. It's difficult to define optimal, because if we say that optimal is the curve that minimizes mistakes in the two sets, then it might not work as well on unseen data, which is known as overfitting. SVMs help avoid overfitting by finding the points in the dataset closest to the line separating the two classes; these are support vectors. Then the SVM will maximize the margin (distance) between the vectors and separation line. After it finds a line using the margins, it will create a formula, called the weight vector, that shows how important each feature is in deciding the classification. Then when you need to classify an object, it will plug its features into the weight vector and get a score, if the score is positive then its one class and negative will be the other. Before this training, it's important that the different features are on a similar scale so the model is not biased by a feature with large numbers. This is done by subtracting the mean and dividing by the standard deviation.

Bernard Ng, Rebecca K. Reh, Sara Mostafavi, A practical guide to applying machine learning to infant EEG data, *Developmental Cognitive Neuroscience*, Volume 54, 2022, 101096, ISSN 1878-9293, <https://doi.org/10.1016/j.dcn.2022.101096>.