Support Vector Machines (SVMs)

1 Support Vector Machines (SVMs)

Classifier learning can be seen as finding a curve that best separates two sets of points. The goal is to find an **optimal curve** out of infinitely many possibilities. However, defining "optimal" is challenging. For example, if we define the optimal curve as the one that minimizes mistakes within the two sets, it may not generalize well to unseen data—a problem known as **overfitting**.

Support Vector Machines (SVMs) address overfitting by focusing on the points in the dataset that are closest to the line separating the two classes; these points are called **support vectors**. SVMs maximize the **margin** (distance) between the support vectors and the separating line. After finding this line using the margins, the SVM creates a formula called the **weight vector**, which indicates the importance of each feature in deciding the classification.

For classification, the features of an object are plugged into the weight vector to get a score:

- A **positive score** indicates one class.
- A **negative score** indicates the other class.

1.1 Feature Scaling

Before training, it's essential to ensure that different features are on a similar scale to prevent bias from features with large numbers. This is typically done by:

1. Subtracting the mean of each feature.
2. Dividing by the standard deviation of each feature.

2 References

- 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