**Notes**

**Support Vector Machine (SVM)** - algorithm that is primarily used for classification and regression tasks. Its main objective is to find the optimal hyperplane that best separates the data into different classes or predicts a continuous target variable.

*Key concepts and functionality of SVMs:*

**Hyperplane:** In the context of SVM, a hyperplane is a decision boundary that separates data into different classes. For binary classification, the hyperplane aims to maximize the margin between the two classes. The margin is defined as the distance between the hyperplane and the nearest data points from each class.

For 2D – line, 3D – plane(плоскость), for nD – is Hyperplane.

**Support Vectors:** Support vectors are the data points that are closest to the hyperplane and are crucial in defining the margin. These support vectors are the most challenging data points to classify and play a significant role in determining the position and orientation of the hyperplane.

A diagram of a support line

Description automatically generated

**Margin:** The margin is the space between the hyperplane and the nearest support vectors. SVM aims to find the hyperplane that maximizes this margin. Maximizing the margin often leads to better generalization on unseen data.

A diagram of a graph

Description automatically generated

**Kernel Trick:** SVMs can handle both linear and nonlinear classification and regression tasks using kernel functions. A kernel function allows the SVM to implicitly map the data into a higher-dimensional space where it can find a linear hyperplane to separate the classes. Common kernel functions include linear, polynomial, radial basis function (RBF), and sigmoid kernels.

A diagram of a graph

Description automatically generated

**Regularization Parameter (C):** As mentioned earlier, the parameter 'C' controls the trade-off between maximizing the margin and minimizing classification errors. It helps balance the bias-variance trade-off in the model.

A diagram of a normalized curve

Description automatically generated with medium confidence

**Soft Margin SVM:** In practical scenarios, data may not be perfectly separable, and there might be outliers. In such cases, you can use a "soft margin" SVM by allowing some data points to be misclassified or fall within the margin. The parameter 'C' controls the degree of tolerance for misclassification.

A diagram of a diagram of a diagram

Description automatically generated with medium confidence

SVMs are widely used in various applications, including image classification, text classification, bioinformatics, and more. They are known for their effectiveness in high-dimensional spaces and their ability to handle complex decision boundaries. However, SVMs can be computationally expensive for large datasets, and tuning hyperparameters like 'C' and the choice of kernel function is important for achieving good performance.

**C** - is the regularization parameter. This parameter controls the trade-off between maximizing the margin and minimizing the classification error. Here's what it means:

**Large C (e.g., C=1e3 or C=1000):** When you set C to a large value, the SVM classifier will try to classify all training examples correctly. This means the algorithm will have a smaller margin but may have a higher classification accuracy on the training data. This could lead to overfitting if the data is noisy or not well-separated.

**Small C (e.g., C=1e-3 or C=0.001):** When you set C to a small value, the SVM classifier will aim for a larger margin, potentially allowing some training examples to be misclassified. This increases the bias of the model but can lead to better generalization to unseen data, which is useful when dealing with noisy or overlapping data.