Interview Questions:

Q1.What is a support vector?
Ans.A support vector is a data point that lies closest to the decision boundary (also called the margin) in a Support Vector Machine (SVM) model. These points are critical because they directly influence the position and orientation of the hyperplane that separates different classes.

SVM tries to find the hyperplane that maximizes the margin between classes, and only the support vectors are used in this optimization. If these points were removed, the position of the decision boundary would change. Thus, support vectors are the most "informative" elements of the dataset for classification.

Q 2. What does the C parameter do?
Ans.The C parameter in machine learning models like Support Vector Machines (SVM) controls the trade-off between achieving a low training error and maintaining a smooth decision boundary. A smaller C value allows the model to have a wider margin by tolerating more misclassifications, promoting better generalization. Conversely, a larger C tries to classify all training points correctly, which can lead to a tighter margin and possible overfitting. Essentially, C regulates how much the model penalizes misclassification during training.

Q 3. What are kernels in SVM?
Ans. Kernels in SVM are functions that transform the input data into a higher-dimensional space, allowing the model to find a linear separating hyperplane even when the data is not linearly separable in the original space. They enable SVMs to capture complex relationships by implicitly mapping data without explicitly computing the transformation. Common kernels include linear, polynomial, and radial basis function (RBF), each suited for different types of data patterns.

Q4. What is the difference between linear and RBF kernel?
Ans. The linear kernel in SVM calculates the similarity between data points using a simple dot product, making it suitable for linearly separable data where a straight line (or hyperplane) can separate classes. It is computationally efficient and works well when the data is already in a form that is easy to separate.

The RBF (Radial Basis Function) kernel, on the other hand, maps data into an infinite-dimensional space using a Gaussian function. It can handle complex, non-linear relationships by creating flexible decision boundaries that adapt to the shape of the data. RBF is powerful for datasets where classes overlap or have intricate patterns that a linear kernel cannot separate.

Q 5.What are the advantages of SVM?
Ans. SVMs are effective in high-dimensional spaces and remain robust when the number of features exceeds the number of samples. They use the kernel trick to model complex, non-linear relationships without explicitly transforming data, making them versatile for various problems. Additionally, SVMs focus on maximizing the margin between classes, which often leads to better generalization and less overfitting compared to other classifiers.

Q 6.Can SVMs be used for regression?
Ans. Yes, SVMs can be used for regression through a variant called Support Vector Regression (SVR). Instead of classifying data points, SVR tries to fit a function within a margin of tolerance to predict continuous values. It uses the same principles as SVM classification, such as maximizing the margin, but focuses on minimizing prediction errors within a specified threshold, making it suitable for regression tasks.

Q7.What happens when data is not linearly separable?

Ans. When data is not linearly separable, a simple straight line or hyperplane cannot perfectly separate the classes. In such cases, SVM uses techniques like the kernel trick to map the data into a higher-dimensional space where a linear separator can be found. Alternatively, it allows some misclassifications by introducing a soft margin controlled by the C parameter, balancing between fitting the data and maintaining a smooth boundary.

Q8. How is overfitting handled in SVM?
Ans. Overfitting in SVM is managed primarily through the C parameter, which controls the trade-off between fitting the training data and keeping the decision boundary smooth. A smaller C allows more misclassifications, encouraging a simpler model that generalizes better. Additionally, choosing an appropriate kernel and tuning its parameters, like gamma in the RBF kernel, helps prevent the model from capturing noise as patterns, thus reducing overfitting.