**1. What is a support vector?**

A **support vector** is a data point that lies closest to the decision boundary (or hyperplane) in an SVM model. These points are critical because they directly affect the position and orientation of the hyperplane. The SVM algorithm tries to maximize the margin (distance) between the hyperplane and these support vectors.

**2. What does the C parameter do?**

The **C parameter** controls the trade-off between achieving a low error on the training data and maintaining a large margin.

* **High C**: Less tolerant to misclassification (low bias, high variance).
* **Low C**: More tolerant to misclassification to allow a larger margin (high bias, low variance).

**3. What are kernels in SVM?**

**Kernels** are functions that transform the data into a higher-dimensional space so that it becomes linearly separable. This allows SVM to handle non-linear relationships without explicitly computing the transformation (thanks to the **kernel trick**).

Common kernels:

* **Linear**
* **Polynomial**
* **RBF (Radial Basis Function)**
* **Sigmoid**

**4. What is the difference between linear and RBF kernel?**

| **Feature** | **Linear Kernel** | **RBF Kernel** |
| --- | --- | --- |
| Transformation | No transformation; works in original space | Maps data to infinite-dimensional space |
| Use case | When data is linearly separable | When data has complex, non-linear relationships |
| Computational cost | Lower | Higher |
| Parameters | Only C | C and γ (gamma) |

**5. What are the advantages of SVM?**

* **Effective in high-dimensional spaces**
* **Robust to overfitting**, especially with a proper kernel and regularization
* **Works well even with few samples** (especially with a clear margin of separation)
* **Flexible through the use of kernels** for non-linear decision boundaries

**6. Can SVMs be used for regression?**

Yes, SVMs can be used for **regression** tasks using **Support Vector Regression (SVR)**. Instead of trying to classify data points, SVR tries to fit as many data points as possible within a margin of tolerance (epsilon).

**7. What happens when data is not linearly separable?**

When data is not linearly separable:

* **Soft-margin SVM** is used (with C parameter) to allow some misclassifications.
* **Kernel trick** is applied to map data to a higher-dimensional space where it becomes linearly separable.

**8. How is overfitting handled in SVM?**

Overfitting is handled in SVM through:

* **Regularization parameter (C)**: Controls the trade-off between margin size and classification error.
* **Kernel choice**: Avoiding overly complex kernels helps prevent overfitting.
* **Gamma (γ) parameter** in RBF kernel: Controls the influence of a single training example. High gamma can lead to overfitting.