1. What is a support vector?

- Support vectors are the **data points that lie closest to the decision boundary (hyperplane)**.

- They are critical because they directly define the position and orientation of the boundary.

- Removing non–support vectors won't change the boundary, but removing a support vector will.

---

2. What does the C parameter do?

- **C is a regularization parameter** in SVM.

- A **large C** $\rightarrow$ tries to classify all training points correctly (low bias, high variance, risk of overfitting).

- A **small C** $\rightarrow$ allows more misclassifications but gives a simpler, smoother boundary (higher bias, lower variance).

---

3. What are kernels in SVM?

- Kernels are mathematical functions that let SVMs work in **higher-dimensional spaces** without explicitly computing the coordinates.

- They measure similarity between data points.

- Common kernels:

  o **Linear**: no transformation, just dot product.

  o **Polynomial**: compares points using polynomial features.

  o **RBF (Gaussian)**: compares points using distance similarity (localized).

---

4. Difference between Linear and RBF kernel?

- **Linear kernel**: assumes data is linearly separable (straight line/plane boundary).

- **RBF kernel**: creates a flexible, non-linear boundary by mapping data into an infinite-dimensional feature space.

- If features are already linearly separable $\rightarrow$ Linear is better.

- If boundaries are complex/curved $\rightarrow$ RBF works better.

---

5. What are the advantages of SVM?

- Works well in **high-dimensional spaces**.

- Effective when number of features > number of samples.

- Robust to overfitting (especially with good regularization).

- Supports both **linear and non-linear classification** via kernels.

- Often performs well even with limited data.

---

6. Can SVMs be used for regression?

- ✅ Yes → called **Support Vector Regression (SVR)**.

- Instead of finding a boundary, SVR fits a function that predicts continuous values within a tolerance margin (ε-insensitive loss).

---

7. What happens when data is not linearly separable?

- SVM introduces a **soft margin** (controlled by C) to allow misclassifications.

- Or uses a **kernel trick** (like RBF) to transform data into higher dimensions where it may become separable.

---

8. How is overfitting handled in SVM?

- By tuning **C** (controls margin softness).

- By adjusting **kernel parameters** (e.g., gamma in RBF).

- By using **cross-validation** to pick hyperparameters.

- With **feature scaling/selection** to avoid noise dominance.