Q1. Random Forest Regressor is a machine learning algorithm that belongs to the ensemble learning family and is used for regression tasks. It is an extension of the decision tree algorithm and operates by constructing a multitude of decision trees during training and outputting the mean prediction of the individual trees as the final prediction.

Q2. Random Forest Regressor reduces the risk of overfitting by:

- Training multiple decision trees on different subsets of the training data (bootstrap samples).
- Randomly selecting a subset of features at each split of the decision tree.
- Aggregating the predictions of multiple trees, which helps in smoothing out the predictions and reducing the variance.
-

Q3. Random Forest Regressor aggregates the predictions of multiple decision trees by averaging the predictions made by individual trees. For regression tasks, the final prediction of the Random Forest Regressor is the mean prediction of all the trees.

Q4. The hyperparameters of Random Forest Regressor include:

- n_estimators: The number of decision trees in the forest.
- max_features: The maximum number of features to consider for splitting a node.
- max_depth: The maximum depth of each decision tree.
- min_samples_split: The minimum number of samples required to split an internal node.
- min_samples_leaf: The minimum number of samples required to be at a leaf node.
-

Q5. The main difference between Random Forest Regressor and Decision Tree Regressor lies in their approach to building the model. While Random Forest Regressor builds multiple decision trees and averages their predictions to reduce overfitting, Decision Tree Regressor builds a single decision tree without any aggregation. Random Forest Regressor tends to have better generalization performance and is less prone to overfitting compared to Decision Tree Regressor.

Q6. Advantages of Random Forest Regressor:

- Reduced risk of overfitting.
- Good performance on both small and large datasets.
- Handles high-dimensional data well.
- Provides estimates of feature importance. Disadvantages:
- Can be computationally expensive, especially with a large number of trees.
- Less interpretable compared to individual decision trees.
- May not perform well with noisy data.
-

Q7. The output of Random Forest Regressor is a continuous numerical value, which represents the predicted target variable for a given set of input features.

Q8. While Random Forest Regressor is primarily designed for regression tasks, Random Forest can also be adapted for classification tasks using the Random Forest Classifier. It operates similarly to the regressor but is tailored for predicting discrete class labels instead of continuous numerical values