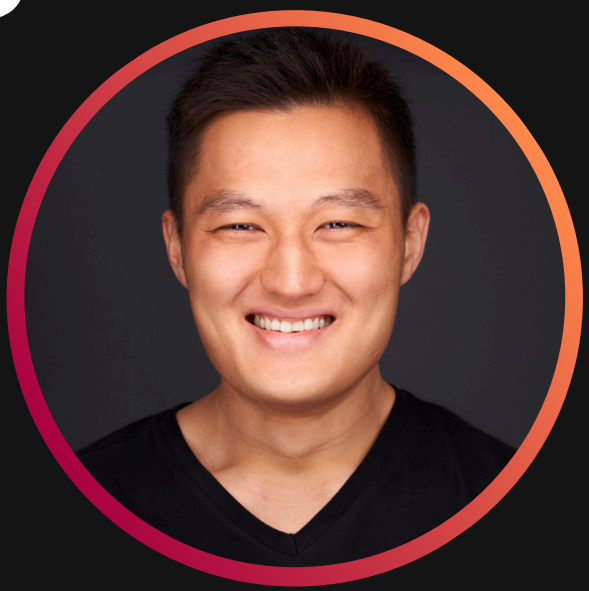
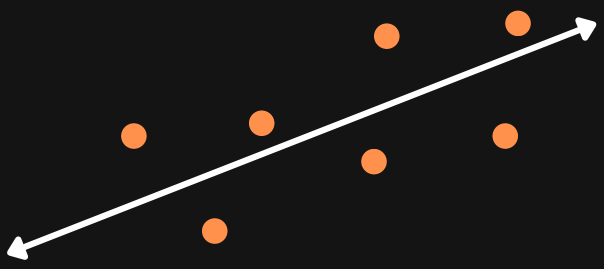
Common ML Algorithms Asked in Interviews



Dan Lee

Linear Regresion

Predicts a continuous outcome by fitting a linear relationship between input features and the target variable.



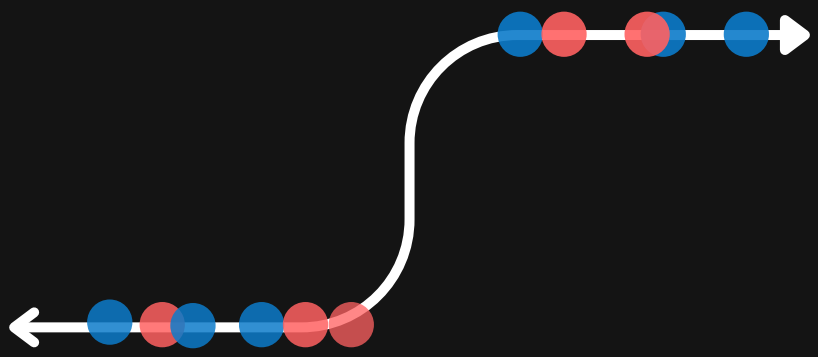
Interview Questions

1. How does linear regression handle multicollinearity, and why is it a problem?
2. What assumptions are made in linear regression, and how can you test for them?
3. How do you interpret the coefficients in a linear regression model?

Next Tip

Logistic Regression

Estimates the probability of a binary outcome by applying a logistic function to a linear combination of input features.



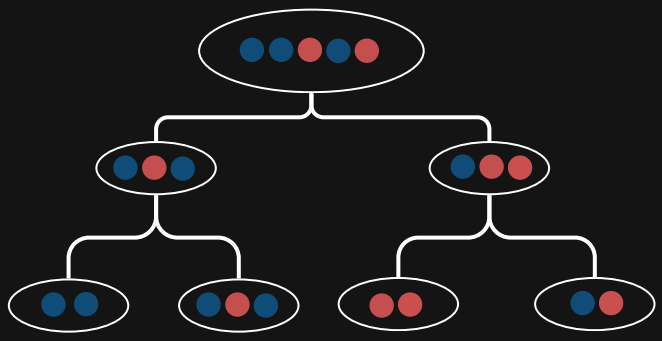
Interview Questions

1. How do you evaluate the performance of a logistic regression model?
2. What is the role of the sigmoid function in logistic regression?
3. What is the difference between L1 and L2 regularization in logistic regression?

Next Tip

Decision Tree

Creates a flowchart-like model to make predictions by recursively splitting data based on feature values.



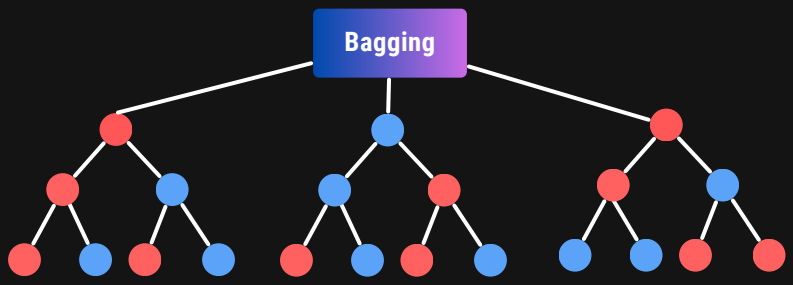
Interview Questions

1. How does a decision tree handle categorical and continuous variables?
2. What are the common metrics used to determine the quality of a split in decision trees?
3. How do you prevent overfitting in decision trees?

Next Tip

Random Forest

Combines multiple decision trees to improve prediction accuracy and reduce overfitting by averaging their outputs.



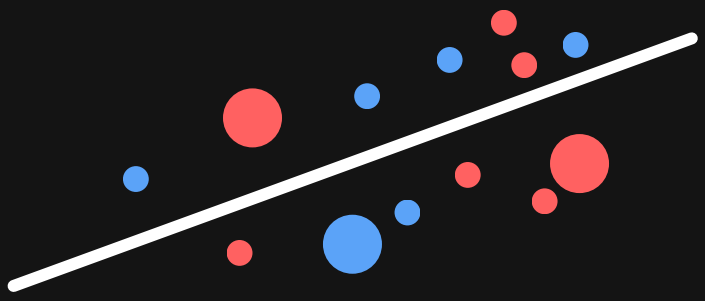
Interview Questions

1. What is the significance of bootstrapping in random forests?
2. How does the random forest algorithm reduce overfitting compared to a single decision tree?
3. What is the impact of increasing the number of trees in a random forest?

Next Tip

Boosted Trees

Builds an ensemble of trees sequentially, optimizing each one to correct errors made by the previous trees.



Interview Questions

1. How does gradient boosting differ from random forests?
2. What is the role of the learning rate in gradient boosting?
3. How does gradient boosting handle overfitting, and what hyperparameters can you tune to address it?
4. Explain how boosting sequentially corrects errors from previous models.

Next Tip

K-Means

Groups data into clusters by minimizing the variance within each cluster based on feature similarity.

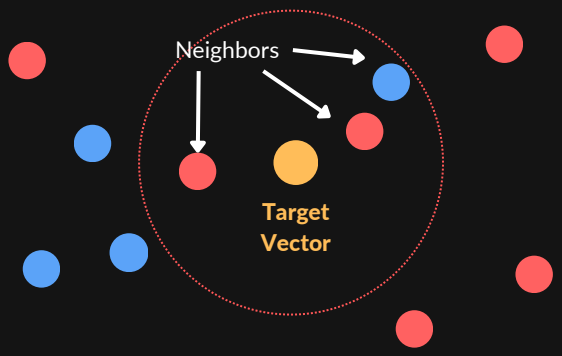
Interview Questions

1. How do you choose the optimal number of clusters in K-Means?
2. What is the role of the centroid in K-Means clustering?
3. How does K-Means handle high-dimensional data, and what are the potential challenges?

Next Tip

K-Nearest Neighbor

Predicts the outcome for a data point by analyzing the majority outcome of its closest neighbors in feature space.



Interview Questions

1. How does the choice of k affect the performance of the KNN algorithm?
2. What are the pros and cons of using Euclidean distance as a distance metric in KNN?
3. How does KNN handle high-dimensional data, and what are common strategies to address challenges?

Next Tip

Like This Post?

✍️ Save

🔁 Repost

👇 Comment