1. Which of the following algorithms is NOT considered a statistical-based algorithm in classification? a) Naive Bayes b) Logistic Regression c) k-Nearest Neighbors d) Decision Tree

Answer: d) Decision Tree

Explanation: Decision tree-based algorithms are not considered statistical-based algorithms. They are tree-like structures where an internal node represents a feature, the branch represents a decision rule, and each leaf node represents the outcome.

---

2. In k-Nearest Neighbors (k-NN) algorithm, how is the class label of a new instance determined? a) By majority voting of its k nearest neighbors b) By fitting a linear regression model c) By maximizing the posterior probability d) By minimizing the Euclidean distance to the centroid

Answer: a) By majority voting of its k nearest neighbors

Explanation: In k-NN, the class label of a new instance is determined by the majority class among its k nearest neighbors in the feature space.

---

3. Which algorithm is known for its ability to handle non-linear decision boundaries effectively? a) Logistic Regression b) Decision Tree c) Support Vector Machine (SVM) d) Random Forest

Answer: b) Decision Tree

Explanation: Decision trees can capture non-linear decision boundaries effectively by partitioning the feature space into rectangular regions.

---

4. What is a key characteristic of Neural Network-based classification algorithms? a) They are interpretable and easily explainable. b) They are based on statistical probability theory. c) They consist of interconnected layers of nodes/neurons. d) They are primarily used for binary classification tasks.

Answer: c) They consist of interconnected layers of nodes/neurons.

Explanation: Neural network-based classification algorithms consist of interconnected layers of nodes/neurons, which allow them to learn complex patterns and relationships in the data.

---

5. Which algorithm is primarily based on a set of IF-THEN rules to classify instances? a) Naive Bayes b) Decision Tree c) k-Nearest Neighbors d) Logistic Regression

Answer: b) Decision Tree

Explanation: Decision tree-based algorithms use a set of IF-THEN rules to classify instances based on the values of input features.

---

6. Which algorithm is considered a probabilistic classifier? a) Decision Tree b) k-Nearest Neighbors c) Logistic Regression d) Support Vector Machine (SVM)

Answer: c) Logistic Regression

Explanation: Logistic Regression is a probabilistic classifier that models the probability of a binary outcome.

---

7. Which of the following algorithms is most sensitive to feature scaling? a) Decision Tree b) Naive Bayes c) k-Nearest Neighbors d) Random Forest

Answer: c) k-Nearest Neighbors

Explanation: k-Nearest Neighbors algorithm relies on distance metrics, such as Euclidean distance, which can be sensitive to feature scaling. Therefore, feature scaling is crucial for k-NN.

---

8. What is the main advantage of Rule-based algorithms in classification tasks? a) They are computationally efficient. b) They can handle missing values well. c) They are highly accurate

for large datasets. d) They provide interpretable classification rules.

Answer: d) They provide interpretable classification rules.

Explanation: Rule-based algorithms produce classification rules that are easily interpretable by humans, aiding in understanding the decision-making process.

---

9. Which algorithm can be prone to overfitting when dealing with noisy data? a) Decision Tree b) Support Vector Machine (SVM) c) Naive Bayes d) Random Forest

Answer: a) Decision Tree

Explanation: Decision trees can easily overfit noisy data by creating overly complex trees that capture noise rather than signal in the data.

---

10. In probabilistic classifiers, how are class probabilities typically estimated? a) By maximizing the likelihood function b) By minimizing the error rate c) By fitting a linear regression model d) By calculating the distance to the nearest neighbors

Answer: a) By maximizing the likelihood function

Explanation: Probabilistic classifiers estimate class probabilities by maximizing the likelihood

function, which represents the probability of observing the training data given the model parameters.

Related posts:

1. Introduction to Information Security
2. Introduction to Information Security MCQ
3. Introduction to Information Security MCQ
4. Symmetric Key Cryptography MCQ
5. Asymmetric Key Cryptography MCQ
6. Authentication & Integrity MCQ
7. E-mail, IP and Web Security MCQ