1

Question: What is the primary objective of the study mentioned in the text?

A) To compare the performance of different machine learning models

B) To predict cholesterol and glucose levels using machine learning

C) To analyze the correlation between cardiovascular disease and lifestyle changes

D) To develop a new algorithm for predicting heart disease

Correct Answer: B

2

Question: Which dataset is used in the study for predicting cardiovascular disease?

A) Kaggle Dataset

B) Cardiovascular Disease Dataset

C) Machine Learning Dataset

D) Healthcare Dataset

Correct Answer: B

3

Question: What is the name of the algorithm that builds a sequence of trees to improve predictions by correcting errors from prior trees?

A) Gradient Boosting Machine (GBM)

B) Random Forest

C) Decision Tree

D) Naive Bayes

Correct Answer: A

4

Question: Which of the following is NOT a type of machine learning model used in the study?

A) K-Nearest Neighbor

B) Logistic Regression

C) Decision Tree

D) Support Vector Machine

Correct Answer: D

5

Question: What is the purpose of the preprocessing step in the study?

A) To handle issues like incomplete data and categorical columns

B) To implement machine learning models

C) To evaluate the performance of models

D) To collect data

Correct Answer: A

6

Question: Which model achieved the highest accuracy in predicting cholesterol levels in the 70:30 split?

A) Gradient Boosting Machine (GBM)

B) Random Forest

C) Logistic Regression

D) Decision Tree

Correct Answer: A

7

Question: What is the name of the technique used to assign a class to a test point based on the majority label at the terminal node?

A) Classification

B) Regression

C) Clustering

D) Decision Tree

Correct Answer: A

8

Question: Which of the following is a performance metric used to evaluate the models in the study?

A) Accuracy

B) Precision

C) Recall

D) All of the above

Correct Answer: D

9

Question: What is the name of the algorithm that creates multiple decision trees and combines their outputs to improve accuracy?

A) Random Forest

B) Gradient Boosting Machine (GBM)

C) Decision Tree

D) Naive Bayes

Correct Answer: A

10

Question: Which model performed the best in predicting glucose levels in the 70:30 split?

A) Gradient Boosting Machine (GBM)

B) Random Forest

C) Logistic Regression

D) Decision Tree

Correct Answer: B

11

Question: What is the purpose of the feature scaling step in the study?

A) To eliminate the mean and adjust to unit variance

B) To handle missing values

C) To implement machine learning models

D) To evaluate the performance of models

Correct Answer: A

12

Question: Which of the following is a type of ensemble method used in the study?

A) Gradient Boosting Machine (GBM)

B) Random Forest

C) Both A and B

D) None of the above

Correct Answer: C

13

Question: What is the name of the algorithm that assumes feature independence and calculates the posterior probability of each class?

A) Naive Bayes

B) Logistic Regression

C) Decision Tree

D) Random Forest

Correct Answer: A

14

Question: Which model achieved the highest accuracy in predicting the combined target of cholesterol and glucose levels?

A) Gradient Boosting Machine (GBM)

B) Random Forest

C) Logistic Regression

D) Decision Tree

Correct Answer: A

15

Question: What is the purpose of the bootstrap sampling step in the Random Forest algorithm?

A) To select a random subset of features

B) To select a random subset of data

C) To implement machine learning models

D) To evaluate the performance of models

Correct Answer: B

16

Question: Which of the following is a type of machine learning model that is sensitive to noise?

A) K-Nearest Neighbor

B) Logistic Regression

C) Decision Tree

D) Random Forest

Correct Answer: A

17

Question: What is the name of the technique used to calculate the probability of each class in the Logistic Regression model?

A) Sigmoid function

B) Decision Tree

C) Random Forest

D) Naive Bayes

Correct Answer: A

18

Question: Which model performed the worst in predicting cholesterol levels in the 70:30 split?

A) Gradient Boosting Machine (GBM)

B) Random Forest

C) Decision Tree

D) Naive Bayes

Correct Answer: C

19

Question: What is the purpose of the stopping criterion in the Decision Tree algorithm?

A) To prevent overfitting

B) To handle missing values

C) To implement machine learning models

D) To evaluate the performance of models

Correct Answer: A

20

Question: Which of the following is a type of performance metric used to evaluate the models in the study?

A) Precision

B) Recall

C) F1 score

D) All of the above

Correct Answer: D

21

Question: What is the name of the algorithm that uses a sigmoid function to predict the probability of each class?

A) Logistic Regression

B) Decision Tree

C) Random Forest

D) Naive Bayes

Correct Answer: A

22

Question: Which model achieved the highest accuracy in predicting glucose levels in the 80:20 split?

A) Gradient Boosting Machine (GBM)

B) Random Forest

C) Logistic Regression

D) Decision Tree

Correct Answer: B

23

Question: What is the purpose of the feature selection step in the Random Forest algorithm?

A) To select a random subset of features

B) To select a random subset of data

C) To implement machine learning models

D) To evaluate the performance of models

Correct Answer: A

24

Question: Which of the following is a type of ensemble method that builds multiple decision trees and combines their outputs?

A) Gradient Boosting Machine (GBM)

B) Random Forest

C) Both A and B

D) None of the above

Correct Answer: C

25

Question: What is the name of the algorithm that uses a CatBoost classification algorithm to predict cardiovascular disease risk?

A) Gradient Boosting Machine (GBM)

B) Random Forest

C) Decision Tree

D) Naive Bayes

Correct Answer: A

26

Question: Which model performed the best in predicting the combined target of cholesterol and glucose levels in the 70:30 split?

A) Gradient Boosting Machine (GBM)

B) Random Forest

C) Logistic Regression

D) Decision Tree

Correct Answer: A

27

Question: What is the purpose of the study mentioned in the text?

A) To develop a new algorithm for predicting heart disease

B) To compare the performance of different machine learning models

C) To predict cholesterol and glucose levels using machine learning

D) To analyze the correlation between cardiovascular disease and lifestyle changes

Correct Answer: C