47. Ques: In data exploration, which statistical measure is used to describe the center of the data distribution?

    a-: Covariance

    b-: Mode

    c-: Mean

    d-: Correlation coefficient

    Ans: c-: Mean

48. Ques: Which data exploration technique is used to identify the spread or dispersion of data points around the mean?

    a-: Box plot

    b-: Histogram

    c-: Scatter plot

    d-: Pie chart

    Ans: a-: Box plot

49. Ques: Which data mining technique is used for classification tasks where the target variable has more than two categories?

    a-: K-Nearest Neighbors (K-NN)

    b-: Naive Bayes Classifier

    c-: Decision Tree

    d-: Logistic Regression

    Ans: c-: Decision Tree

50. Ques: What is the main advantage of using the K-Nearest Neighbors (K-NN) algorithm in data mining?

    a-: It works well with high-dimensional data.

    b-: It can handle missing values in the dataset.

    c-: It is easy to interpret the results.

    d-: It adapts well to nonlinear relationships between variables.

    Ans: d-: It adapts well to nonlinear relationships between variables.

51. Ques: Which evaluation metric is used to assess the performance of a regression model in predicting continuous values?

    a-: F1 score

    b-: R-squared (R2)

    c-: Accuracy

d-: Precision

Ans: b-: R-squared (R2)

52. Ques: Which step in the data mining process involves selecting the appropriate data mining algorithm for the task?

a-: Data cleaning

b-: Data transformation

c-: Data modeling

d-: Data exploration

Ans: c-: Data modeling

53. Ques: What is the primary purpose of logistic regression in data mining?

a-: To predict numerical values for a continuous outcome variable.

b-: To classify data into two or more discrete categories.

c-: To identify associations between variables.

d-: To segment the customer base for targeted marketing.

Ans: b-: To classify data into two or more discrete categories.

54. Ques: Which data mining technique is used for unsupervised learning to group similar data points together?

a-: Decision Tree

b-: K-Nearest Neighbors (K-NN)

c-: Logistic Regression

d-: Cluster Analysis

Ans: d-: Cluster Analysis

55. Ques: In multiple linear regression, how many independent variables can be used to predict the dependent variable?

a-: One

b-: Two

c-: Three

d-: Any number

Ans: d-: Any number

56. Ques: Which data mining technique is used for classification tasks where the target variable has more than two categories?

a-: K-Nearest Neighbors (K-NN)

b-: Naive Bayes Classifier

c-: Decision Tree

d-: Logistic Regression

Ans: c-: Decision Tree

57. Ques: What is the primary advantage of using the Decision Tree algorithm for classification tasks?

a-: It can handle missing values in the dataset.

b-: It works well with high-dimensional data.

c-: It is easy to interpret and visualize the results.

d-: It can capture nonlinear relationships between variables.

Ans: c-: It is easy to interpret and visualize the results.

58. Ques: In the context of model building, what does the term "overfitting" mean?

a-: The model is too complex and performs well on training data but poorly on unseen data.

b-: The model is too simple and does not capture the underlying patterns in the data.

c-: The model is well-balanced and performs equally well on training and testing data.

d-: The model is unable to make predictions due to missing data.

Ans: a-: The model is too complex and performs well on training data but poorly on unseen data.

59. Ques: Which evaluation metric is used to assess the performance of a classification model?

a-: Mean Absolute Error (MAE)

b-: Root Mean Squared Error (RMSE)

c-: F1 score

d-: R-squared (R2)

Ans: c-: F1 score

60. Ques: What is the primary purpose of logistic regression in data mining?

a-: To predict numerical values for a continuous outcome variable.

b-: To classify data into two or more discrete categories.

c-: To identify associations between variables.

d-: To segment the customer base for targeted marketing.

Ans: b-: To classify data into two or more discrete categories.

61. Ques: Which data mining technique is used for unsupervised learning to group similar data points together?

    a-: Decision Tree

    b-: K-Nearest Neighbors (K-NN)

    c-: Logistic Regression

    d-: Cluster Analysis

    Ans: d-: Cluster Analysis

62. Ques: In multiple linear regression, how many independent variables can be used to predict the dependent variable?

    a-: One

    b-: Two

    c-: Three

    d-: Any number

    Ans: d-: Any number

63. Ques: Which data mining technique is used for classification tasks where the target variable has more than two categories?

    a-: K-Nearest Neighbors (K-NN)

    b-: Naive Bayes Classifier

    c-: Decision Tree

    d-: Logistic Regression

    Ans: c-: Decision Tree

64. Ques: What is the primary advantage of using the Decision Tree algorithm for classification tasks?

    a-: It can handle missing values in the dataset.

    b-: It works well with high-dimensional data.

    c-: It is easy to interpret and visualize the results.

    d-: It can capture nonlinear relationships between variables.

    Ans: c-: It is easy to interpret and visualize the results.

65. Ques: In the context of model building, what does the term "overfitting" mean?

    a-: The model is too complex and performs well on training data but poorly on unseen data.

b-: The model is too simple and does not capture the underlying patterns in the data.

c-: The model is well-balanced and performs equally well on training and testing data.

d-: The model is unable to make predictions due to missing data.

Ans: a-: The model is too complex and performs well on training data but poorly on unseen data.

66. Ques: Which evaluation metric is used to assess the performance of a classification model?

a-: Mean Absolute Error (MAE)

b-: Root Mean Squared Error (RMSE)

c-: F1 score

d-: R-squared (R2)

Ans: c-: F1 score

67. Ques: What is the primary purpose of logistic regression in data mining?

a-: To predict numerical values for a continuous outcome variable.

b-: To classify data into two or more

 discrete categories.

c-: To identify associations between variables.

d-: To segment the customer base for targeted marketing.

Ans: b-: To classify data into two or more discrete categories.

68. Ques: Which data mining technique is used for unsupervised learning to group similar data points together?

a-: Decision Tree

b-: K-Nearest Neighbors (K-NN)

c-: Logistic Regression

d-: Cluster Analysis

Ans: d-: Cluster Analysis

69. Ques: In multiple linear regression, how many independent variables can be used to predict the dependent variable?

a-: One

b-: Two

c-: Three

d-: Any number

Ans: d-: Any number

70. Ques: Which data mining technique is used for classification tasks where the target variable has more than two categories?

a-: K-Nearest Neighbors (K-NN)

b-: Naive Bayes Classifier

c-: Decision Tree

d-: Logistic Regression

Ans: c-: Decision Tree

71. Ques: What is the primary advantage of using the Decision Tree algorithm for classification tasks?

a-: It can handle missing values in the dataset.

b-: It works well with high-dimensional data.

c-: It is easy to interpret and visualize the results.

d-: It can capture nonlinear relationships between variables.

Ans: c-: It is easy to interpret and visualize the results.

72. Ques: In the context of model building, what does the term "overfitting" mean?

a-: The model is too complex and performs well on training data but poorly on unseen data.

b-: The model is too simple and does not capture the underlying patterns in the data.

c-: The model is well-balanced and performs equally well on training and testing data.

d-: The model is unable to make predictions due to missing data.

Ans: a-: The model is too complex and performs well on training data but poorly on unseen data.

73. Ques: Which evaluation metric is used to assess the performance of a classification model?

a-: Mean Absolute Error (MAE)

b-: Root Mean Squared Error (RMSE)

c-: F1 score

d-: R-squared (R2)

Ans: c-: F1 score

74. Ques: What is the primary purpose of logistic regression in data mining?

a-: To predict numerical values for a continuous outcome variable.

b-: To classify data into two or more discrete categories.

c-: To identify associations between variables.

d-: To segment the customer base for targeted marketing.

Ans: b-: To classify data into two or more discrete categories.

75. Ques: Which data mining technique is used for unsupervised learning to group similar data points together?

a-: Decision Tree

b-: K-Nearest Neighbors (K-NN)

c-: Logistic Regression

d-: Cluster Analysis

Ans: d-: Cluster Analysis

76. Ques: In multiple linear regression, how many independent variables can be used to predict the dependent variable?

a-: One

b-: Two

c-: Three

d-: Any number

Ans: d-: Any number

77. Ques: Which data mining technique is used for classification tasks where the target variable has more than two categories?

a-: K-Nearest Neighbors (K-NN)

b-: Naive Bayes Classifier

c-: Decision Tree

d-: Logistic Regression

Ans: c-: Decision Tree

8. Ques: What is the primary advantage of using the Decision Tree algorithm for classification tasks?

a-: It can handle missing values in the dataset.

b-: It works well with high-dimensional data.

c-: It is easy to interpret and visualize the results.

d-: It can capture nonlinear relationships between variables.

Ans: c-: It is easy to interpret and visualize the results.

79. Ques: In the context of model building, what does the term "overfitting" mean?

a-: The model is too complex and performs well on training data but poorly on unseen data.

b-: The model is too simple and does not capture the underlying patterns in the data.

c-: The model is well-balanced and performs equally well on training and testing data.

d-: The model is unable to make predictions due to missing data.

Ans: a-: The model is too complex and performs well on training data but poorly on unseen data.

80. Ques: Which evaluation metric is used to assess the performance of a classification model?

a-: Mean Absolute Error (MAE)

b-: Root Mean Squared Error (RMSE)

c-: F1 score

d-: R-squared (R2)

Ans: c-: F1 score

81. Ques: What is the primary purpose of logistic regression in data mining?

a-: To predict numerical values for a continuous outcome variable.

b-: To classify data into two or more discrete categories.

c-: To identify associations between variables.

d-: To segment the customer base for targeted marketing.

Ans: b-: To classify data into two or more discrete categories.

82. Ques: Which data mining technique is used for unsupervised learning to group similar data points together?

a-: Decision Tree

b-: K-Nearest Neighbors (K-NN)

c-: Logistic Regression

d-: Cluster Analysis

Ans: d-: Cluster Analysis

83. Ques: In multiple linear regression, how many independent variables can be used to predict the dependent variable?

a-: One

b-: Two

c-: Three

d-: Any number

Ans: d-: Any number

84. Ques: Which data mining technique is used for classification tasks where the target variable has more than two categories?

a-: K-Nearest Neighbors (K-NN)

b-: Naive Bayes Classifier

c-: Decision Tree

d-: Logistic Regression

Ans: c-: Decision Tree

85. Ques: What is the primary advantage of using the Decision Tree algorithm for classification tasks?

a-: It can handle missing values in the dataset.

b-: It works well with high-dimensional data.

c-: It is easy to interpret and visualize the results.

d-: It can capture nonlinear relationships between variables.

Ans: c-: It is easy to interpret and visualize the results.

86. Ques: In the context of model building, what does the term "overfitting" mean?

a-: The model is too complex and performs well on training data but poorly on unseen data.

b-: The model is too simple and does not capture the underlying patterns in the data.

c-: The model is well-balanced and performs equally well on training and testing data.

d-: The model is unable to make predictions due to missing data.

Ans: a-: The model is too complex and performs well on training data but poorly on unseen data.

87. Ques: Which evaluation metric is used to assess the performance of a classification model?

a-: Mean Absolute Error (MAE)

b-: Root Mean Squared Error (RMSE)

c-: F1 score

d-: R-squared (R2)

Ans: c-: F1 score

88. Ques: What is the primary purpose of logistic regression

in data mining?

    a-: To predict numerical values for a continuous outcome variable.

    b-: To classify data into two or more discrete categories.

    c-: To identify associations between variables.

    d-: To segment the customer base for targeted marketing.

    Ans: b-: To classify data into two or more discrete categories.

89. Ques: Which data mining technique is used for unsupervised learning to group similar data points together?

    a-: Decision Tree

    b-: K-Nearest Neighbors (K-NN)

    c-: Logistic Regression

    d-: Cluster Analysis

    Ans: d-: Cluster Analysis

90. Ques: In multiple linear regression, how many independent variables can be used to predict the dependent variable?

    a-: One

    b-: Two

    c-: Three

    d-: Any number

    Ans: d-: Any number

91. Ques: Which data mining technique is used for classification tasks where the target variable has more than two categories?

    a-: K-Nearest Neighbors (K-NN)

    b-: Naive Bayes Classifier

    c-: Decision Tree

    d-: Logistic Regression

    Ans: c-: Decision Tree