Section – A(20 questions)

**1. What is the primary goal of Artificial Intelligence (AI)?**

a) To mimic human intelligence b) To automate repetitive tasks

c) To make machines smarter than humans d) To solve complex mathematical problems

**2. Which of the following is a supervised learning algorithm?**

a) K-means clustering b) Decision tree c) Apriori algorithm d) K-nearest neighbors (KNN)

**3. What is the main advantage of using Python for AI and ML development?**

a) High speed execution b) Strongly typed language

c) Large community support and libraries d) Easy memory management

**4. Which library is commonly used for numerical operations in Python?**

a) TensorFlow b) Scikit-learn c) Pandas d) NumPy

**5. What does EDA stand for in the context of data analysis?**

a) Effective Data Analysis b) Exploratory Data Augmentation

c) Exploratory Data Analysis d) Efficient Data Aggregation

**6. Which library in Python is primarily used for data visualization?**

a) NumPy b) TensorFlow c) Matplotlib d) Scikit-learn

**7. What is the purpose of feature engineering in machine learning?**

a) To extract information from unstructured data

b) To improve model performance by creating new features

c) To visualize data using graphs and charts

d) To preprocess data before model training

**8. What is the process of converting text data into numerical form called?**

a) Tokenizationb) Normalizationc) Vectorizationd) Dimensionality reduction

**9. Which of the following is an example of unsupervised learning algorithm?**

a) Linear regression b) K-means clustering c) Decision tree d) Support Vector Machine (SVM)

**10. What does SQL stand for in the context of data handling?**

a) Structured Query Language b) Statistical Query Language

c) Sequential Query Logic d) Simplified Query Layout

**11. What is the primary purpose of regularization in machine learning?**

a) To increase model complexity b) To reduce model complexity

c) To remove outliers from the data d) To enhance model interpretability

**12. Which statistical measure represents the central tendency of a dataset?**

a) Mean b) Median c) Mode d) Range

**13. In Pandas, which method is used to drop missing values from a DataFrame?**

a) `remove\_na()` b) `dropna()` c) `drop\_missing()` d) `clean()`

**14. Which plot is used to visualize the distribution of a single numerical variable?**

a) Scatter plot b) Bar plot c) Histogram d) Box plot

**15. What does the acronym API stand for in the context of web development?**

a) Advanced Programming Interface b) Application Protocol Interface

c) Automated Program Invocation d) Application Programming Interface

**16. Which method is used to split a dataset into training and testing sets in scikit-learn?**

a) `split\_dataset()` b) `train\_test\_split()` c) `divide\_data()` d) `prepare\_data()`

**17. Which of the following is NOT a classification algorithm?**

a) Decision tree b) K-means clustering c) Logistic regression d) Support Vector Machine (SVM)

**18. What is the primary objective of normalization in data preprocessing?**

a) To reduce the range of values b) To make data distribution normal

c) To scale features to a similar range d) To remove outliers from the data

**19. Which library is commonly used for deep learning applications in Python?**

a) Matplotlib b) Pandas c) TensorFlow d) Scikit-learn

**20. What is the purpose of cross-validation in machine learning?**

a) To increase model complexity b) To reduce overfitting

c) To preprocess data d) To visualize data

Section – B(15 questions)

**21. Given the dataset: `[10, 15, 20, 25, 30]`, what is the range of the data?**

a) 20 b) 15 c) 25 d) 30

**22. Consider a DataFrame with the columns 'Sales' and 'Expenses'. If the total sales revenue is $5000 and the total expenses are $3000, what is the profit margin (profit as a percentage of sales)?**

a) 40% b) 30% c) 20% d) 10%

**23. Calculate the coefficient of variation (CV) for the dataset: `[12, 15, 18, 22, 25]`.**

a) 16% b) 14% c) 12% d) 10%

**24. Given a DataFrame with the columns 'Length' and 'Width', if the average length is 10 meters and the average width is 5 meters, what is the average area?**

a) 15 square metersb) 25 square metersc) 50 square metersd) 75 square meters

**25. Calculate the mean absolute deviation (MAD) for the dataset: `[5, 8, 12, 15, 20]`.**

a) 4.8 b) 4.0 c) 3.2 d) 2.4

26. **Given a DataFrame with the columns 'Temperature' (in Celsius) and 'Humidity' (in percentage). If the average temperature is 25°C and the average humidity is 60%, what is the relative humidity in degrees Celsius?**

a) 15°C b) 18°C c) 20°C d) 22.5°C

**27. Calculate the median absolute deviation (MAD) for the dataset: `[10, 15, 20, 25, 30]`.**

a) 5.6 b) 4.8 c) 3.6 d) 2.4

28. **Given a DataFrame with the columns 'Population' and 'Area'. If the average population is 5000 and the average area is 100 square kilometers, what is the population density in individuals per square kilometer?**

a) 20 b) 50 c) 100 d) 200

29. **Calculate the root mean squared error (RMSE) for the dataset: `[10, 15, 20, 25, 30]`, if the predicted values are `[12, 18, 21, 24, 29]`.**

a) 1.58 b) 1.73 c) 2.08 d) 2.24

**30. Given a DataFrame with the columns 'Length' (in meters) and 'Width' (in meters). If the total area is 500 square meters and the total perimeter is 100 meters, what is the average length?**

a) 10 meters b) 15 meters c) 20 meters d) 25 meters

**31. Calculate the sum of squared errors (SSE) for the dataset: `[8, 10, 12, 15, 20]`, if the mean is 13.**

a) 30 b) 35 c) 40 d) 45

**32. Given a DataFrame with the columns 'Price' and 'Quantity'. If the average price is $10 and the average quantity sold is 50, what is the total revenue?**

a) $250 b) $500 c) $750 d) $1000

**33. Calculate the mean for the dataset: `[15, 18, 20, 22, 25]`.**

a) 20 b) 18 c) 22 d) 25

**34. Given a DataFrame with the columns 'Speed' (in km/h) and 'Time' (in hours). If the average speed is 60 km/h and the total time is 5 hours, what is the total distance traveled?**

a) 250 km b) 300 km c) 350 km d) 400 km

**35. Calculate the standard deviation for the dataset: `[5, 8, 12, 15, 20]`.**

a) 3.5 b) 4.5 c) 5.5 d) 6.5