قائمة الاسئلة

امتحان نهاية الفصل الدراسي الثاني - للعام الجامعي 1446 هـ - الموافق -2025/2024مـكلية الحاسوب وتكنولوجيا المعلومات :: اختياري عل

د.مروة عبدالرحمن الهادي

(1 1. Which of the following is the main step in the Data Science Workflow?
(1 Apply clustering and classification algorithms -
(2 Define the business problem +
(3 Skip exploratory data analysis -
(4 Directly deploy without validation -

(2 Which of the following is an example of a primary data source?
(1 News articles -
(2 Surveys +
(3 Online blogs -
(4 Social media posts -

(3 Which type of memory in Arduino Uno retains its values even after the board is powered off?
(1 Flash memory -
(2 SRAM -
(3 EEPROM +
(4 ROM -

(4 What is MQTT primarily used for in IoT?
(1 High-speed gaming -
(2 Lightweight messaging protocol for IoT devices +
(3 Large-scale database management -
(4 Hardware device programming -

(5 Which of the following is NOT one of the original 3 Vs of Big Data?
(1 Volume -
(2 Velocity -
(3 Veracity +
(4 Variety -

(6 What is the main issue with imbalanced datasets in machine learning?
(1 Overfitting on the training data -
(2 Poor performance on the minority class +
(3 High variance in predictions -
(4 Low accuracy on the majority class -

(7 What is the default threshold for probabilistic output in most classification models?
(1 0.25 -
(2 0.5 +
(3 0.75 -
(4 1 -

(8 Which metric is best suited for evaluating performance on imbalanced datasets?
(1 Accuracy -
(2 Mean squared error -
(3 ROC-AUC +
(4 R-squared -

(9 What is the primary objective of feature selection in machine learning?
(1 To increase the number of features -
(2 To reduce classification error +
(3 To make the model more complex -
(4 To decrease the dataset size -

(10 Which of the following is NOT a type of feature selection method?

Filter methods - (1
Wrapper methods - (2
Embedded methods - (3
Clustering methods + (4

What is the purpose of using a variance threshold in feature selection? (11
To eliminate features with low variance + (1
To rank features by importance - (2
To optimize model hyperparameters - (3

What is Deep Learning? (12
A subset of Machine Learning that uses mathematical functions to map input to output + (1
A replacement for Artificial Intelligence - (2
A method exclusively used for image processing - (3

What are the three layers in a Deep Neural Network? (13
Input layer, Hidden layers, Output layer + (1
Neural layer, Training layer, Testing layer - (2
Preprocessing layer, Activation layer, Output layer - (3
Input layer, Processing layer, Output layer - (4

Which neural network type is best suited for sequential data like text or audio? (14
Convolutional Neural Network (CNN) - (1
Recurrent Neural Network (RNN) + (2
Fully Connected Network (FCN) - (3

What is the purpose of Exploratory Data Analysis (EDA)? (15
Detecting patterns and guiding further analysis + (1
Deploying solutions - (2
Encoding categorical data - (3

What is the primary goal of Data Science? (16
Building advanced machine learning models - (1
Extracting knowledge from data and having a new product + (2
Competing in machine learning competitions - (3
Storing large amounts of data - (4

What is NOT an example of structured data? (17
XML files with schema - (1
Comma-separated values (CSV) - (2
Website videos + (3
Relational database tables - (4

2. What is the key distinction between Data Mining and Data Science? (18
Data Mining focuses on hardware integration, while Data Science focuses on coding - (1
Data Science includes Data Mining and also covers cleaning, visualization, and deployment + (2
Data Mining is broader than Data science - (3

Which of the following is an automated method of data collection using sensors? (19
Conducting surveys - (1
Observations - (2
IoT device integration + (3
Filling spreadsheets manually - (4

What is a common issue in data quality? (20
Overloaded datasets - (1
Unlabeled axes in visualizations - (2
Noisy or irrelevant data + (3
Using too many APIs - (4

What are the three key characteristics of Big Data according to Gartner?* (21
Volume, Velocity, Variety + (1
Volume, Veracity, Value - (2
Variety, Visualization, Validity - (3
Variability, Vulnerability, Velocity - (4
What is the main goal of Data Science?* (22
To create machine learning models - (1
To extract knowledge and insights from data + (2
To visualize large datasets - (3
To improve computational speed - (4
Which of the following is an example of an unstructured data type?* (23
CSV files - (1
XML files - (2
Videos + (3
Clickstream data - (4
In CRISP-DM, what is the first phase of the data science process?* (24
Data Understanding - (1
Modeling - (2
Business Understanding + (3
Deployment - (4
What is the role of data visualization in Data Science?* (25
To collect data from sensors - (1
To represent insights graphically + (2
To clean the dataset - (3
To analyze algorithm performance - (4
What is an essential skill for data scientists to work collaboratively?* (26
Knowledge of databases - (1
Programming skills - (2
Communication skills + (3
Expertise in machine learning algorithms - (4
What is an example of a potential application for data science in the year 2050?* (27
Manual data collection in healthcare - (1
Predictive modeling in agriculture + (2
Analysis of static spreadsheets - (3
Real-time decision-making in social media - (4
What is the main purpose of IoT in data science?* (28
To store large amounts of data - (1
To generate real-time data for predictive modeling and decision-making + (2
To replace traditional data collection methods - (3
To design hardware for data storage - (4
Which of the following methods is an example of an embedded method for feature selection? (29
Recursive Feature Elimination (RFE) - (1
Information Gain - (2
LASSO (L1 Regularization) + (3
Why is class imbalance a problem in machine learning? (30
It increases dataset size - (1
It improves model accuracy - (2
It causes bias toward the majority class, leading to poor performance on the minority class + (3
It reduces computational cost - (4

Which method is NOT a solution for class imbalance? (31
Cost-Sensitive Learning - (1
Oversampling - (2
Undersampling - (3
Data Normalization + (4
What does the AUC-ROC curve measure? (32
The correlation between independent features - (1
The error rate in regression models - (2
The trade-off between True Positive Rate and False Positive Rate + (3
The relationship between precision and recall - (4
What is the main goal of feature engineering? (33
To decrease dataset dimensionality - (1
To create new features that improve model performance + (2
To increase the dataset size - (3
To remove all categorical features - (4
Why is feature scaling important in machine learning? (34
It increases computational complexity - (1
It reduces dataset size - (2
It improves the performance of algorithms + (3
It eliminates missing values - (4
What does One-Hot Encoding do? (35
Detects missing values - (1
Converts categorical features into multiple binary variables + (2
Standardizes numerical data - (3
Removes duplicate features - (4
Why is feature selection important? (36
It removes categorical data - (1
It helps remove irrelevant and redundant features + (2
It increases the number of features - (3
It decreases model accuracy - (4
What is the purpose of feature extraction in machine learning? (37
To replace numerical values with categorical labels - (1
To create new meaningful features from raw data + (2
To remove all missing values - (3
To generate duplicate features - (4
What does the Box Plot visualization help with in data preprocessing? (38
Scaling numerical features - (1
Encoding categorical variables - (2
Identifying outliers + (3
Replacing missing values - (4
Which data preprocessing technique is most suitable for dealing with highly numerical data? (39
One-hot encoding - (1
Min-Max Scaling - (2
Standardization - (3
Log transformation + (4
What is the primary goal of data balancing in machine learning? (40
To remove missing values from the dataset - (1
To increase dataset size - (2
To make the model favor the majority class - (3

(4 To ensure equal representation of classes for fair model learning +