Harun Abi'den Gelen

1. What is Artificial Intelligence (AI)?
a) A programming language
b) Systems that mimic human intelligence
c) An operating system
d) A data storage technique
Answer: b

2. Which of the following is not a type of AI?
a) Narrow AI
b) Superintelligent AI
c) Generative AI
d) Artificial Machine
Answer: d

3. What milestone did IBM's Deep Blue achieve in 1997?
a) Passed the Turing Test
b) Defeated a world chess champion
c) Launched the first chatbot
d) Invented machine learning
Answer: b

4. What is the primary goal of machine learning (ML)?
a) To replace humans in all tasks
b) To learn patterns from data
c) To create expert systems
d) To design neural networks
Answer: b

5. What is supervised learning?
a) Learning with labeled data
b) Learning without any guidance
c) Learning from rewards and punishments
d) Learning through neural networks only
Answer: a

6. Which algorithm is best suited for classification tasks?
a) Linear Regression
b) Logistic Regression
c) K-Means Clustering
d) Gradient Descent
Answer: b

7. In reinforcement learning, what drives the agent's learning?
a) Data labeling
b) Rewards and penalties
c) Feature engineering
d) Neural networks
Answer: b

8. Which type of data is considered structured?
a) Images
b) Spreadsheets
c) Text documents
d) Audio files
Answer: b) Spreadsheets

9. What does data anonymization ensure?
a) Data accuracy
b) Privacy of sensitive information
c) Higher computation speeds
d) Enhanced data visualization
Answer: b) Privacy of sensitive information

10. What is a recommender system used for?
a) Predicting numerical values
b) Suggesting items based on preferences
c) Classifying images
d) Reducing data dimensionality
Answer: b

11. What is the primary function of the sigmoid function in logistic regression?
a) Normalize data
b) Map outputs to probabilities between 0 and 1
c) Extract features
d) Improve accuracy
Answer: b

12. What is the main advantage of a Random Forest algorithm?
a) Low memory usage
b) Generates a single decision tree
c) Reduces overfitting by combining multiple trees
d) Only works for regression tasks
Answer: c

13. What is PCA commonly used for?
a) Data labeling
b) Dimensionality reduction
c) Supervised learning
d) Data wrangling
Answer: b) Dimensionality reduction

14. Which is a type of unsupervised learning?
a) Regression
b) Classification
c) Clustering
d) Decision Trees
Answer: c) Clustering

15. What is a key feature of Generative Adversarial Networks (GANs)?
a) Supervised learning
b) Generating synthetic data
c) Data encryption
d) Optimizing loss functions
Answer: b) Generating synthetic data

16. What is a major challenge of using SVMs?
a) Requires large datasets
b) Inefficient in memory usage
c) Computationally intensive for large data
d) Works poorly with high-dimensional data
Answer: c

17. What is the primary goal of data preparation?
a) To split data into training and testing sets
b) To transform raw data into a usable format
c) To visualize trends in the data
d) To perform dimensionality reduction
Answer: b

18. Which metric is used to evaluate classification models?
a) Mean Squared Error
b) Accuracy
c) Root Mean Square Error
d) R-squared
Answer: b) Accuracy

19. What is the main assumption in Naive Bayes classifiers?
a) Data points are dependent
b) Features are conditionally independent
c) Data is always normalized
d) Data contains outliers
Answer: b

20. What is overfitting?
a) A model failing to capture data trends
b) A model being too simple
c) A model performing well on training data but poorly on new data
d) A model optimizing cost too quickly
Answer: c

21. What is the primary objective of feature scaling?
a) Handle missing data
b) Convert categorical variables
c) Standardize feature ranges
d) Improve model accuracy
Answer: c) Standardize feature ranges

22. Which algorithm is suitable for predicting continuous values?
a) Naive Bayes
b) Decision Tree Regression
c) K-Means Clustering
d) Random Forest Classification
Answer: b) Decision Tree Regression

23. What is the kernel trick in SVMs used for?
a) Handling outliers
b) Transforming data into higher dimensions
c) Normalizing data
d) Improving training speeds
Answer: b

24. What does a soft margin in SVM allow?
a) Strict separation of data points
b) Perfect accuracy in all cases
c) Tolerance for misclassifications
d) High computational efficiency
Answer: c) Tolerance for misclassifications

25. What is the first step in building a machine learning model?
a) Train the model
b) Collect data
c) Evaluate the model
d) Split the data
Answer: b) Collect data

26. Which of the following best describes Artificial Intelligence (AI)?
A) A field focused on performing mathematical calculations
B) Machines mimicking human-like intelligence to perform tasks
C) A system that uses only rule-based logic
D) A type of robotic hardware
Answer: B

27. What was a major milestone in AI in 1997?
A) Turing's "Computing Machinery and Intelligence"
B) Dartmouth Conference
C) Deep Blue defeating Garry Kasparov
D) The release of ChatGPT
Answer: C

28. What is the primary purpose of Machine Learning (ML)?
A) Automate repetitive tasks
B) Teach machines to learn from data
C) Simulate human consciousness
D) Create rule-based systems
Answer: B

29. Which of the following is NOT a type of ML?
A) Supervised Learning
B) Unsupervised Learning
C) Randomized Learning
D) Reinforcement Learning
Answer: C

30. What type of AI is capable of performing general intellectual tasks like a human?
A) Narrow AI
B) General AI
C) Superintelligent AI
D) Reinforcement AI
Answer: B

31. What algorithm is most suitable for predicting continuous outcomes?
A) Logistic Regression
B) Linear Regression
C) K-Nearest Neighbors
D) Decision Trees
Answer: B

32. Which ML type uses labeled data for training?
A) Unsupervised Learning
B) Reinforcement Learning
C) Supervised Learning
D) Generative Learning
Answer: C

33. Which of these is a classification algorithm?
A) Polynomial Regression
B) Logistic Regression
C) Linear Regression
D) PCA
Answer: B

34. What is the primary goal of reinforcement learning?
A) Group similar data points
B) Learn patterns from labeled data
C) Maximize cumulative rewards over time
D) Reduce dimensionality of datasets
Answer: C

35. What technique does SVM use to handle non-linear data?
A) Gradient Descent
B) Kernel Trick
C) Clustering
D) Normalization
Answer: B

36. Which of the following is a feature of data anonymization?
A) Adding noise to data
B) Labeling data points
C) Grouping similar data
D) Predicting class labels
Answer: A

37. Which type of regression is used for binary classification?
A) Linear Regression
B) Logistic Regression
C) Polynomial Regression
D) Decision Tree Regression
Answer: B

38. What is PCA primarily used for?
A) Classification
B) Regression
C) Dimensionality Reduction
D) Reinforcement Learning
Answer: C

39. What is an example of unsupervised learning?
A) Linear Regression
B) Decision Trees
C) K-Means Clustering
D) Naive Bayes
Answer: C

40. In recommender systems, collaborative filtering suggests items based on:
A) Past purchases of the same user
B) Similarities between items
C) Preferences of other users with similar tastes
D) Random selection of items
Answer: C

41. What does ETL stand for in data warehousing?
A) Extract, Transform, Load
B) Evaluate, Test, Learn
C) Explore, Train, Label
D) Extract, Train, Learn
Answer: A

42. What does "overfitting" refer to in machine learning?
A) A model that performs well on unseen data
B) A model that performs poorly on training data
C) A model that performs well on training data but poorly on unseen data
D) A model with high bias
Answer: C

43. What is a common cost function used in linear regression?
A) Cross-Entropy Loss
B) Mean Squared Error
C) Gini Index
D) Information Gain
Answer: B

44. What does "feature engineering" involve?
A) Normalizing data
B) Creating new features from existing data
C) Reducing the dimensionality of data
D) Splitting data into training and testing sets
Answer: B

45. Which of these is NOT a step in building a machine learning model?
A) Data Collection
B) Algorithm Design
C) Data Labeling
D) Model Deployment
Answer: B

46. What is the purpose of the sigmoid function in logistic regression?
A) Minimize errors
B) Normalize input data
C) Map predictions to probabilities
D) Find clusters in the dataset
Answer: C

47. Which algorithm uses Bayes' Theorem for classification?
A) Decision Trees
B) Support Vector Machines
C) Naive Bayes
D) K-Nearest Neighbors
Answer: C

48. Which supervised learning algorithm is ensemble-based?
A) Linear Regression
B) Random Forest
C) K-Means
D) PCA
Answer: B

49. Which ML algorithm uses Q-Learning?
A) Reinforcement Learning
B) Supervised Learning
C) Dimensionality Reduction
D) Clustering
Answer: A

50. What type of AI performs only narrow tasks?
A) Narrow AI
B) General AI
C) Super AI
D) Symbolic AI
Answer: A

51. Which AI milestone involved the creation of the first natural language processing program?
A) Dartmouth Conference
B) ELIZA Program
C) AlphaGo vs. Lee Sedol
D) IBM's Deep Blue
Answer: B

52. What is a key feature of Narrow AI?
A) Ability to perform all human intellectual tasks
B) Focus on specific tasks or domains
C) Self-awareness and consciousness
D) General problem-solving skills
Answer: B

53. Which of the following is an example of a supervised learning task?
A) Clustering customers based on behavior
B) Predicting house prices
C) Reducing data dimensionality
D) Generating synthetic data
Answer: B

54. What technique reduces the number of features in a dataset while retaining important information?
A) Feature Engineering
B) Dimensionality Reduction
C) Data Balancing
D) Clustering
Answer: B

55. What is the purpose of data splitting in machine learning?
A) To clean the data
B) To transform categorical variables
C) To create training, validation, and test sets
D) To balance imbalanced datasets
Answer: C

56. Which type of AI is still largely theoretical?
A) Narrow AI
B) General AI
C) Superintelligent AI
D) Symbolic AI
Answer: C

57. What method is used to group data points into clusters in unsupervised learning?
A) Decision Trees
B) K-Means
C) Logistic Regression
D) Random Forest
Answer: B

58. What does overfitting indicate in a model?
A) The model is performing well on test data
B) The model memorizes training data but performs poorly on unseen data
C) The model generalizes well to new data
D) The model has no errors
Answer: B

59. Which machine learning algorithm is based on Bayes' Theorem?
A) Linear Regression
B) Naive Bayes
C) K-Means Clustering
D) Decision Tree
Answer: B

60. What is reinforcement learning primarily used for?
A) Classification tasks
B) Predicting probabilities
C) Learning optimal actions through rewards
D) Dimensionality reduction
Answer: C

61. What is a major advantage of Random Forest over Decision Trees?
A) It handles only binary classification
B) It reduces overfitting by combining multiple trees
C) It uses gradient descent optimization
D) It requires less computational power
Answer: B

62. Which step in ML involves creating new variables from existing ones?
A) Data Cleaning
B) Feature Engineering
C) Data Balancing
D) Normalization
Answer: B

63. What does ETL stand for?
A) Explore, Train, Learn
B) Extract, Transform, Load
C) Evaluate, Test, Label
D) Execute, Transform, Learn
Answer: B

64. What is the purpose of Principal Component Analysis (PCA)?
A) Prediction of continuous values
B) Feature scaling
C) Dimensionality reduction
D) Generating synthetic data
Answer: C

65. Which algorithm is most suitable for predicting customer churn?
A) K-Means Clustering
B) Logistic Regression
C) PCA
D) Naive Bayes
Answer: B

66. What is the output of clustering algorithms?
A) Predictions
B) Groups or clusters
C) Probabilities
D) Numerical values
Answer: B

67. What does a confusion matrix evaluate?
A) Model's speed
B) Data quality
C) Classification performance
D) Regression accuracy
Answer: C

68. Which algorithm finds the hyperplane that separates classes?
A) Logistic Regression
B) Support Vector Machines
C) K-Means Clustering
D) Decision Trees
Answer: B

69. Which regression type is suitable for non-linear relationships?
A) Linear Regression
B) Polynomial Regression
C) Logistic Regression
D) Support Vector Regression
Answer: B

70. What does gradient descent optimize?
A) Learning rate
B) Model architecture
C) Cost function
D) Data balance
Answer: C

71. Which model is best suited for text classification tasks?
A) Naive Bayes
B) Linear Regression
C) K-Means
D) Random Forest
Answer: A

72. What is a key characteristic of symbolic AI?
A) Learning from large datasets
B) Use of symbols and logic for problem-solving
C) Predicting continuous values
D) Clustering similar data points
Answer: B

73. What does feature scaling achieve in machine learning?
A) Improves model interpretability
B) Reduces dimensionality
C) Normalizes feature ranges for better performance
D) Removes irrelevant features
Answer: C

74. What is the role of synthetic data in ML?
A) Enhance model interpretability
B) Reduce overfitting
C) Replace real-world data for privacy reasons
D) Reduce data redundancy
Answer: C

75. Which evaluation metric combines precision and recall?
A) F1-Score
B) Accuracy
C) Mean Absolute Error
D) AUC-ROC
Answer: A

TEORİ SORULARI

1. Discuss the main differences between clustering and classification in machine learning.

- **Clustering**: An **unsupervised learning** technique that groups similar data points into clusters based on inherent patterns. No predefined labels are used.
- **Classification**: A **supervised learning** technique where the model learns to map input data to predefined labels (e.g., "spam" or "not spam").
- Key Difference: Classification requires labeled data, while clustering works with unlabeled data.

2. Explain the steps involved in data cleaning and their significance.

- **Handling missing values**: Filling, dropping, or imputing missing data.
- **Removing duplicates**: Ensuring data uniqueness to avoid redundancy.
- **Fixing structural errors**: Correcting typos, inconsistent formatting, etc.
- **Handling outliers**: Detecting and treating anomalous values.
- **Standardizing data**: Ensuring consistent units and formats.
- **Significance**: Improves data quality, which directly impacts model accuracy and reliability.

3. How does PCA help in handling high-dimensional datasets?

- PCA reduces dimensions by transforming features into a smaller set of uncorrelated components, preserving maximum variance.
- Benefits: Reduces computational complexity, avoids the curse of dimensionality, and improves visualization and interpretation.

4. What is the importance of balancing datasets in classification tasks?

- Prevents bias towards the majority class.
- Ensures that the model learns meaningful patterns for all classes.
- Balancing techniques: Oversampling (e.g., SMOTE), undersampling, and using class-weighted loss functions.

5. Describe the advantages of Random Forest over single Decision Trees.
Reduced overfitting due to ensemble averaging.

- Better generalization and stability.
- Can handle missing data and is less sensitive to noisy data.

6. How does reinforcement learning differ from supervised and unsupervised learning?

- Reinforcement Learning: Learns through interactions with an environment, optimizing rewards.
- Supervised Learning: Trained on labeled data.
- Unsupervised Learning: Finds patterns in unlabeled data.

7. Explain the use of the sigmoid function in logistic regression.

- Converts linear outputs into probabilities between 0 and 1.
- Helps in binary classification by mapping predictions to class probabilities.

8. What are the strengths and weaknesses of Naive Bayes classifiers?

- **Strengths**: Simple, fast, works well with small datasets, effective for text classification.
- **Weaknesses**: Assumes feature independence, which may not hold true in all datasets.

9. Describe the process of hyperparameter tuning in machine learning.

- Process: Selecting the best hyperparameters (e.g., learning rate, tree depth) using techniques like grid search, random search, or Bayesian optimization.
- Importance: Enhances model performance by finding optimal configurations.

10. What is overfitting, and how can it be prevented in machine learning models? How does overfitting occur, and what methods can be used to prevent it?

Overfitting:

Overfitting happens when a model learns the training data too well, including noise and irrelevant patterns, making it perform poorly on new, unseen data.

Causes:

- **Too complex model** (e.g., deep networks, overgrown decision trees).
- **Small or noisy training data**.
- **Too many irrelevant features**.

Prevention:

1. Use **cross-validation**.
2. Apply **regularization** (L1, L2 penalties).
3. Stop training early (**early stopping**).
4. Simplify the model (reduce features or parameters).
5. Use **dropout** for neural networks.
6. Gather more training data or augment it.
7. Use **ensemble methods** (e.g., Random Forest).

11. Explain the difference between supervised, unsupervised, and reinforcement learning.

- **Supervised**: Labeled data, learns mapping from input to output.
- **Unsupervised**: Unlabeled data, finds patterns (e.g., clustering).
- **Reinforcement**: Learns by maximizing rewards through interaction.

12. Describe the steps to build a machine learning model.

1. Define the problem.
2. Collect and prepare data.
3. Explore and visualize data.
4. Feature engineering and selection.
5. Split data into training and testing sets.
6. Choose and train a model.
7. Evaluate and optimize.
8. Deploy and monitor.

13. What is the significance of data preparation in machine learning?

- Ensures data quality and relevance.
- Improves model accuracy and reliability by eliminating noise, errors, and inconsistencies.

14. Compare linear regression and logistic regression.

- **Linear Regression**: Predicts continuous outcomes.
- **Logistic Regression**: Predicts probabilities for binary or categorical outcomes.
- Key Difference: Output type (continuous vs. categorical).

15. Discuss the advantages and limitations of support vector machines.

- **Advantages**: Effective for high-dimensional data, works well with non-linear boundaries using the kernel trick.
- **Limitations**: Computationally expensive for large datasets, sensitive to choice of kernel and parameters.

16. What is feature engineering, and why is it important in machine learning?

- Process of creating, transforming, or selecting features to improve model performance.
- Importance: Helps the model better understand data, leading to improved predictions.

17. How do recommender systems function, and what are the types of filtering they use?

- Function: Suggests items to users based on preferences or behavior.
- Types:
  - Collaborative Filtering: Based on user-item interactions.
  - Content-Based Filtering: Based on item characteristics.
  - Hybrid Filtering: Combines both.

18. Explain the role of the kernel trick in SVM and how it helps classify non-linear data.

- Maps data to higher dimensions where it becomes linearly separable.
- Enables SVM to classify non-linear data using kernels like RBF or polynomial.

19. What are the differences between data anonymization and data labeling?

- **Data Anonymization**: Removing or masking identifiers to protect privacy.
- **Data Labeling**: Assigning meaningful tags or labels to data for supervised learning.