Advanced Concepts in Artificial Intelligence

Artificial Intelligence · Answer Key · 18 Questions

1. Which paper introduced the Transformer architecture, which serves as the basis for most modern Large Language Models?

A) Attention Is All You Need
B) Deep Residual Learning for Image Recognition
C) Generative Adversarial Nets
D) Mastering the Game of Go with Deep Neural Networks

2. Who is credited with inventing the term 'Artificial Intelligence' at the 1956 Dartmouth Summer Research Project?

A) Alan Turing
B) John McCarthy
C) Marvin Minsky
D) Claude Shannon

3. In the context of backpropagation, what is the 'vanishing gradient' problem?

A) The loss of data packets during training
B) Gradients becoming too small to update weights effectively
C) The overfitting of the model to noise
D) The hardware overheating during computation

4. Which algorithm is defined as a heuristic search that finds the shortest path by combining the cost to reach a node and the estimated cost to the goal?

A) Breadth-First Search
B) A* search algorithm
C) Dijkstra's algorithm
D) Monte Carlo Tree Search

5. What does the 'ReLU' activation function stand for in deep learning?

A) Rectified Linear Unit
B) Recursive Logical Unit
C) Radial Error Linear Update
D) Refined Logic Unit

6. Which specific technique is used to prevent overfitting by randomly setting a fraction of input units to 0 at each update during training?

A) Bagging
B) Dropout
C) Boosting
D) Pruning

7. What is the primary objective of a 'GAN' (Generative Adversarial Network) architecture?

A) To translate languages using recurrent cells
B) To have two neural networks compete against each other
C) To classify images into thousands of categories
D) To map high-dimensional data to lower dimensions

8. In reinforcement learning, what does 'SARSA' stand for?

A) State-Action-Reward-State-Action
B) Search-Algorithm-Random-State-Analysis
C) Statistical-Artificial-Reinforcement-System-Approach
D) State-Action-Regression-Sequential-Analysis

9. Which cognitive architecture was developed by John R. Anderson to model human memory and learning?

A) SOAR
B) ACT-R
C) Cyc
D) OpenCog

10. What is the 'No Free Lunch Theorem' in machine learning?

A) No single model performs best on every possible problem
B) Models cannot be trained without labeled data
C) Computing power is always limited
D) Training time is inversely proportional to accuracy

11. What mathematical operation is the fundamental building block of Convolutional Neural Networks (CNNs)?

A) Cross-correlation
B) Cross-entropy
C) Softmax transformation
D) Backpropagation

12. What is the primary limitation of a perceptron that was famously highlighted by Minsky and Papert in 1969?

A) Inability to handle non-linear activation functions
B) Inability to solve the XOR logical problem
C) High computational complexity
D) Requirement for exponential training data

13. Which company released the 'AlphaGo' system that defeated Lee Sedol in the game of Go?

A) OpenAI
B) Google DeepMind
C) Microsoft Research
D) Facebook AI Research

14. In natural language processing, what are 'word embeddings' like Word2Vec primarily designed to capture?

A) Grammatical syntax rules
B) Semantic relationships and vector similarity
C) The phonetics of spoken language
D) The frequency of character occurrence

15. What is the function of the 'Softmax' layer in a neural network?

A) To normalize output into a probability distribution
B) To compress data into smaller vectors
C) To calculate the derivative of the loss function
D) To initialize weights randomly

16. Which of these is a widely used benchmark dataset for handwritten digit recognition?

A) ImageNet
B) MNIST
C) CIFAR-10
D) COCO

17. What is the 'Transformer' model's 'Self-Attention' mechanism designed to do?

A) Weight the importance of different words in a sequence
B) Reduce the number of parameters in a network
C) Force the network to focus on the start of a sentence
D) Increase the depth of a neural network

18. In deep learning, what is a 'Hyperparameter'?

A) A weight updated during training
B) A parameter whose value is set before the learning process
C) The final output of the model
D) A node in the hidden layer