Started on	Monday, 1 December 2025, 12:48 PM
State	Finished
Completed on	Monday, 1 December 2025, 12:56 PM
Time taken	8 mins 55 secs
Marks	21.00/25.00
Grade	**84.00** out of 100.00

Question 1

Complete

Mark 0.00 out of 1.00

A decision tree tends to overfit when:

- a. The number of classes increases
- b. Gini impurity is used
- c. The dataset is small
- d. The tree depth is very large

Question 2

Complete

Mark 1.00 out of 1.00

A large learning rate usually causes:

- a. Faster convergence
- b. Reduced training time without issues
- c. Increased accuracy
- d. Overshooting and divergence

Question 3

Complete

Mark 0.00 out of 1.00

A perceptron fails when data is:

- a. Non-linearly separable
- b. Centered around zero
- c. High-dimensional
- d. Linearly separable

Question 4

Complete

Mark 1.00 out of 1.00

Batch size affects training primarily by influencing:

- a. Size of the convolution kernel
- b. Activation functions
- c. Gradient noise and stability
- d. Number of hidden layers

Question 5

Complete

Mark 1.00 out of 1.00

In NLP, the purpose of positional encoding in Transformers is to:

- a. Inject sequence order information
- b. Add syntactic structure
- c. Improve tokenization
- d. Reduce vocabulary size

Question 6

Complete

Mark 1.00 out of 1.00

In reinforcement learning, an episode ends when:

- a. A terminal state is reached
- b. Policy becomes random
- c. Reward becomes zero
- d. Learning rate becomes zero

Question 7

Complete

Mark 0.00 out of 1.00

In reinforcement learning, the value function represents:

- a. Opponent action
- b. Immediate reward
- c. Difference between predicted and actual reward
- d. Future expected return from a state

Question 8

Complete

Mark 1.00 out of 1.00

K-Means clustering optimizes which objective?

- a. Maximum likelihood
- b. Classification error
- c. Gradient norm
- d. Sum of squared distances within clusters

Question 9

Complete

Mark 1.00 out of 1.00

L2 regularization penalizes:

- a. Large squared weights
- b. Large absolute weights
- c. Input features
- d. Number of layers

Question 10

Complete

Mark 1.00 out of 1.00

Latent space in an autoencoder represents:

- a. Compressed intermediate representation of data
- b. Loss over epochs
- c. Error gradient
- d. The final predictions

Question 11

Complete

Mark 1.00 out of 1.00

RNNs mainly struggle with long-term dependencies due to:

- a. Low memory
- b. Overfitting
- c. Vanishing and exploding gradients
- d. Slow inference

Question 12

Complete

Mark 1.00 out of 1.00

The key mechanism in Transformers that replaced recurrence is:

- a. Self-attention
- b. Skip connections
- c. Autoencoders
- d. Max pooling

Question 13

Complete

Mark 1.00 out of 1.00

The output of the softmax layer is best described as:

- a. A residual connection
- b. Binary classification result
- c. Normalized probability distribution
- d. One-hot encoded vector

Question 14

Complete

Mark 1.00 out of 1.00

The primary reason CNNs outperform fully connected layers on images is:

- a. They can artificially enlarge data
- b. They train faster
- c. They use more parameters
- d. They exploit spatial locality via filters

Question 15

Complete

Mark 1.00 out of 1.00

The purpose of residual connections in deep networks is to:

- a. Perform pooling
- b. Reduce the number of parameters
- c. Prevent vanishing gradients and improve training
- d. Increase model depth without issues

Question 16

Complete

Mark 1.00 out of 1.00

The universal approximation theorem states that:

- a. A neural network can approximate any continuous function
- b. More neurons always improve accuracy
- c. Neural networks can memorize any dataset
- d. Deep networks always outperform shallow networks

Question 17

Complete

Mark 1.00 out of 1.00

What does tokenization do in NLP?

- a. Removes stop words
- b. Splits text into smaller units (words/subwords)
- c. Converts text to lower case
- d. Parses grammar

Question 18

Complete

Mark 1.00 out of 1.00

Which activation function can output negative values?

- a. Tanh
- b. Softmax
- c. ReLU
- d. Sigmoid

Question 19

Complete

Mark 0.00 out of 1.00

Which ML technique is most suitable for detecting outliers?

- a. PCA
- b. Isolation Forest
- c. SVM (linear)
- d. Linear Regression

Question 20

Complete

Mark 1.00 out of 1.00

Which of the following algorithms assumes features are conditionally independent?

- a. KNN
- b. Logistic Regression
- c. Random Forest
- d. Naive Bayes

Question 21

Complete

Mark 1.00 out of 1.00

Which of the following is a non-linear kernel for SVM?

- a. Polynomial
- b. Hard margin
- c. Linear
- d. Logistic

Question 22

Complete

Mark 1.00 out of 1.00

Which of the following is NOT a feature scaling method?

- a. Standardization
- b. Log scaling
- c. Label encoding
- d. Min–max normalization

Question 23

Complete

Mark 1.00 out of 1.00

Which optimization algorithm adapts the learning rate differently for each parameter?

- a. Adam
- b. Momentum
- c. SGD
- d. SVM

Question 24

Complete

Mark 1.00 out of 1.00

Which type of neural network layer reduces the spatial size of feature maps?

- a. Max Pooling layer
- b. Recurrent layer
- c. Convolutional layer
- d. Fully connected layer

Question 25

Complete

Mark 1.00 out of 1.00

Word2Vec's Skip-Gram model predicts:

- a. Center word from context
- b. Document embeddings
- c. Context words from center word
- d. Both simultaneously