Important Questions

--> Module 1

\* Which type of learning involves an agent learning to interact with an environment to maximize a reward?

- Unsupervised learning

- Supervised learning

- Semi-supervised learning

- Reinforcement learning

\* Which of the following is an example of reinforcement learning?

- A chatbot responding to customer queries

- A robot learning to navigate a maze

- Classifying emails as spam or not spam

- Grouping similar documents together

--> Module 2

\* Which algorithm is commonly used for collaborative filtering in recommendation systems?

- Apriori Algorithm

- Decision Tree

- Matrix Factorization

- K-Means Clustering

\* Which technique is commonly used to convert categorical data into vector representations?

- Linear regression

- One-hot encoding

- Gradient descent

- Singular Value Decomposition

\* What is the main advantage of using matrix factorization for recommendation systems?

- It requires a large amount of labeled data

- It provides perfect predictions

- It can handle sparse data effectively

- It is the simplest algorithm to implement

\* What does the dot product between two word vectors represent in the context of word embeddings?

- The length of the vectors

- The difference between the vectors

- The measure of similarity between the two words

- The angle between the two vectors

--> Module 3

\* Which type of problem is best suited for using a Support Vector Machine (SVM)?

- Dimensionality Reduction

- Classification

- Clustering

- Regression

\* What is the purpose of using a kernel in Support Vector Machines?

- To handle non-linear data

- To normalize the data

- To increase the training data size

- To reduce the number of features

\* What does the term "generalization" mean in the context of machine learning?

- The process of increasing the model complexity

- The process of simplifying a model

- The ability of a model to perform well on unseen data

- The ability of a model to perform well on training data

\* Which of the following methods is used to increase the generalization of a machine learning model?

- Regularization

- Cross-validation

- Increasing the number of features

- Increasing the size of the training set

--> Module 4

\* How are the weights in an ANN typically updated during training?

- By setting them to fixed values

- By randomly adjusting them after each epoch

- By using the error calculated from the loss function

- By using the values from the input data

\* What is backpropagation in the context of ANN?

- A method for initializing weights

- An algorithm used to update the weights based on the error

- A way to add new layers to the network

- A technique for forward propagation of input data

\* What is the purpose of the loss function in an ANN?

- To initialize the weights

- To measure the difference between the predicted output and the actual output

- To determine the structure of the network

- To decide the activation function to be used

\* What does the term "epoch" refer to in the training of an ANN?

- The number of neurons in each layer

- A single update of the weights

- A complete pass through the entire training dataset

- The number of layers in the network

--> Module 5

\* What is the role of the pooling layer in a CNN?

- To apply activation functions

- To increase the number of parameters in the model

- To reduce the spatial dimensions of the input volume

- To compute the dot product between the weights and input

\* Which type of pooling is most commonly used in CNNs?

- Sum pooling

- Average pooling

- Min pooling

- Max pooling

\* What is the "vanishing gradient problem" commonly associated with in deep learning?

- Training deep neural networks

- CNNs

- Fully connected layers

- Shallow neural networks

\* Which of the following is true about the cell state in an LSTM?

- It is updated at each time step and carries long-term information

- It is the same as the hidden state

- It is used to compute the final output of the network

- It does not change over time

\* In an autoencoder, what is the "bottleneck" layer?

- The input layer

- The output layer

- The layer where the data is reconstructed

- The layer with the smallest number of neurons

--> Module 6

\* What is lemmatization in NLP?

- The process of converting words to their root form

- The process of converting all words to lowercase

- The process of removing punctuation from text

- The process of tokenizing the text

\* What does the "Term Frequency" (TF) component of TF-IDF measure?

- The importance of a word in the entire dataset

- The rarity of a word across all documents

- The length of the document

- The frequency of a term in a document

\* What does the term "word embedding" refer to in vector representation?

- A vector that represents an entire document

- A vector that represents the frequency of words

- A vector that represents a single word

- A matrix that represents a set of sentences

\* What is the main innovation introduced by the Transformer model in NLP?

- Deep belief networks for better feature learning

- Attention mechanisms that allow models to focus on different parts of the input

- Convolutional layers for feature extraction

- Recurrent connections for processing sequences

\* In the Transformer model, what is the purpose of the multi-head attention mechanism?

- To increase the depth of the network

- To apply multiple attention mechanisms in parallel to capture different types of relationships

- To reduce the computational cost of the attention mechanism

- To normalize the attention scores

\* What is a key difference between BERT (Bidirectional Encoder Representations from Transformers) and the original Transformer architecture?

- BERT relies on convolutional layers for processing text

- BERT is used only for image processing tasks

- BERT uses only the encoder part of the Transformer

- BERT uses recurrent layers instead of attention layers