1. **Explain One-Hot Encoding**

**Answer:-**

One-hot encoding is a way of representing categorical data as binary vectors. Each unique category is assigned a unique binary vector, where the value at the index of the category is 1 and all other values are 0. For example, if there are 3 categories, then each category would be represented by a 3-dimensional binary vector.

1. **Explain Bag of Words**

**Answer:-**

The bag-of-words model is a way of representing text as a set of words. The model ignores the order of the words and only considers the presence or absence of each word in the text. This is done by creating a vocabulary of all the unique words in the text, and then creating a binary vector for each document that indicates whether or not each word in the vocabulary appears in the document.

1. **Explain Bag of N-Grams**

**Answer:-**

The bag-of-n-grams model is a generalization of the bag-of-words model. Instead of only considering single words, the bag-of-n-grams model considers n-grams, which are sequences of n words. This allows the model to capture more information about the order of words in the text.

1. **Explain TF-IDF**

**Answer:-**

TF-IDF stands for Term Frequency-Inverse Document Frequency. It is a statistical measure that is used to quantify the importance of a word in a document. TF-IDF is calculated as the product of the term frequency (TF) and the inverse document frequency (IDF). The TF is the number of times a word appears in a document, and the IDF is the inverse of the number of documents that contain the word.

1. **What is OOV problem?**

**Answer:-**

The OOV problem (out-of-vocabulary problem) is a problem that occurs when a word does not appear in the vocabulary of a model. This can happen when the model is trained on a dataset of text that does not contain all possible words. When an OOV word is encountered, the model has no way to represent it, and this can lead to errors in the model's predictions.

1. **What are word embeddings?**

**Answer:-**

Word embeddings are a type of distributed representation of words. This means that each word is represented as a vector of real numbers. The vector's dimensions represent the semantic and syntactic properties of the word. Word embeddings are often used in natural language processing tasks such as text classification, machine translation, and question answering.

1. **Explain Continuous bag of words (CBOW)**

**Answer:-**

CBOW (continuous bag of words) is a neural network model that is used to learn word embeddings. CBOW predicts the current word in a sentence given the surrounding words. The model is trained using a supervised learning algorithm.

1. **ExplainSkipGram**

**Answer:-**

SkipGram is a neural network model that is also used to learn word embeddings. SkipGram predicts the surrounding words in a sentence given the current word. The model is trained using a supervised learning algorithm.

1. **Explain Glove Embeddings.**

**Answer:-**

Glove embeddings are a type of word embedding that are trained on a large corpus of text. The corpus is first tokenized and then the co-occurrence statistics of the tokens are calculated. The Glove model is then trained to predict the probability of one token appearing in the context of another token.