# Natural Language Processing (NLP) - Questions and Answers

78. What is natural language processing, and what are its applications?

Natural Language Processing (NLP) is a field of artificial intelligence that focuses on the interaction between computers and human language. It enables machines to understand, interpret, and generate human language. Applications include chatbots, machine translation, sentiment analysis, speech recognition, and text summarization.

79. Describe the steps involved in text preprocessing for NLP.

Text preprocessing involves several steps: tokenization (splitting text into words or sentences), lowercasing, removing punctuation and special characters, removing stop words, stemming or lemmatization, and converting text into numerical representations such as bag-of-words or TF-IDF.

80. What is tokenization, and why is it necessary in NLP?

Tokenization is the process of breaking down text into smaller units called tokens (e.g., words or sentences). It is a crucial step in NLP because it allows algorithms to analyze and process text data in manageable pieces.

81. Explain the concept of stop words in NLP.

Stop words are common words (such as 'the', 'is', 'in') that are often removed from text during preprocessing because they carry little meaningful information and can introduce noise into the analysis.

82. What are n-grams, and how are they used in text analysis?

N-grams are contiguous sequences of 'n' items (typically words) from a given text. They are used in text analysis to capture context and word co-occurrence patterns. For example, bigrams (n=2) and trigrams (n=3) are commonly used in language modeling and text classification.

83. What is sentiment analysis, and how is it performed using NLP techniques?

Sentiment analysis is the process of determining the emotional tone behind a body of text. It is performed using NLP techniques such as text preprocessing, feature extraction, and classification algorithms to categorize text as positive, negative, or neutral.

84. What is named entity recognition (NER) in NLP?

Named Entity Recognition (NER) is an NLP technique used to identify and classify named entities in text into predefined categories such as person names, organizations, locations, dates, and more. It helps in extracting structured information from unstructured text.