



NLP Interview Questions Pack

Overview

This resource brings together the **most relevant NLP interview questions**, grouped by topic and difficulty. It's designed to strengthen both **conceptual clarity** and **practical understanding** — from tokenization to transformers.

Structure

Each question focuses on:

- **Core Idea** — what the concept means.
 - **Application** — how it's used in real NLP tasks.
 - **Example** — when needed, a short illustration.
-

Topics Covered

◆ Text Preprocessing & Representation

1. What is tokenization? Explain different tokenization techniques.
 2. Difference between stemming and lemmatization.
 3. Explain stop-word removal and its importance.
 4. What are n-grams, and how are they used?
 5. How do TF-IDF and CountVectorizer differ?
 6. What are word embeddings?
 7. Explain one-hot encoding and its limitations.
 8. What is subword tokenization (e.g., BPE, WordPiece)?
-

◆ Classical NLP Techniques

1. What is Bag of Words (BoW)?
 2. Explain cosine similarity and its role in text analysis.
 3. What is POS tagging?
 4. Difference between Named Entity Recognition (NER) and POS tagging.
 5. Explain parsing and its types (constituency vs dependency).
 6. What are stop-words and how do they impact text classification?
 7. What is TF-IDF weighting formula?
-

◆ Deep Learning for NLP

1. Explain RNNs, LSTMs, and GRUs in NLP.
 2. What problem does attention solve in sequence models?
 3. Describe the architecture of the Transformer model.
 4. What is positional encoding in Transformers?
 5. Explain the difference between Encoder-only (BERT) and Decoder-only (GPT) architectures.
 6. What is fine-tuning in NLP models?
 7. What is transfer learning in NLP?
-

◆ Word Embeddings & Semantic Understanding

1. What are Word2Vec and GloVe?
 2. Difference between CBOW and Skip-Gram.
 3. What is contextual embedding?
 4. Why are static embeddings limited compared to contextual embeddings?
 5. How does BERT generate contextual word representations?
 6. Explain cosine similarity's role in measuring semantic closeness.
-

◆ Evaluation & Applications

1. What are precision, recall, and F1-score in text classification?
 2. How do you evaluate a sentiment analysis model?
 3. What is BLEU score? How is it used in machine translation?
 4. What is ROUGE score?
 5. How is perplexity used in evaluating language models?
 6. What are common challenges in NLP (ambiguity, sarcasm, OOV)?
 7. Explain transfer learning vs domain adaptation in NLP.
-

◆ Advanced Topics

1. What is prompt engineering in LLMs?
 2. Explain masked language modeling.
 3. What is next-token prediction?
 4. How do retrieval-augmented generation (RAG) systems work?
 5. Explain how sentence transformers differ from BERT.
 6. What is knowledge distillation in NLP?
 7. How does zero-shot or few-shot learning work?
 8. What are large language models (LLMs) and how are they trained?
 9. What are attention heads and why are they important?
 10. How is inference optimization done for large Transformer models?
-

Takeaway

This pack helps learners and professionals **revisit essential NLP theory and applied intuition** — useful for:

- Interview preparation
 - Quick revision before project work
 - Teaching & mentoring resources
-