LLM & NLP Interview Questions – Basics

Brahim Ouhammou

Profile

Hello, my name is **Brahim Ouhammou**. I am a student in **Artificial Intelligence and Machine Learning** at the Faculty of Sciences, and I am also currently studying at **1337 Coding School**.

I have a strong foundation in **programming and problem solving**, with skills in **Python, C, and C++**, as well as experience in **machine learning and deep learning**. I have worked on practical projects involving **CNN-based models, AI-driven systems, and Docker-based infrastructures**.

I also implemented **FT_IRC (Internet Relay Chat)**, an IRC server in **C++** that handles multiple clients simultaneously, focusing on **network programming and socket communication**, as well as **other projects in low-level programming**.

As part of my training at **1337**, I am seeking a **6-month internship**, where I can apply my skills, learn from real-world projects, and continue developing as an **AI and software engineer**.

1 NLP Basics

1. What is NLP?

Natural Language Processing (NLP) is a field of Artificial Intelligence that enables computers to understand, interpret, and generate human language.

2. Common NLP Tasks

- Tokenization
- Part-of-Speech (POS) tagging
- Named Entity Recognition (NER)
- Sentiment Analysis
- Machine Translation
- Text Summarization
- Question Answering

3. What is Tokenization?

Tokenization is the process of splitting text into smaller units such as words, subwords, or characters.

4. Stemming vs Lemmatization

- **Stemming**: Removes word endings crudely (may not produce real words).
- **Lemmatization**: Converts words to their dictionary form using linguistic rules.

5. What is a Corpus?

A corpus is a large collection of text data used for training or evaluating NLP models.

2 Text Representation

6. Bag of Words (BoW)

Bag of Words represents text as word frequency vectors, ignoring word order and grammar.

7. TF-IDF

TF-IDF (Term Frequency–Inverse Document Frequency) measures the importance of a word in a document relative to a corpus.

8. Word Embeddings

Word embeddings are dense vector representations of words that capture semantic meaning.

9. Examples of Embedding Models

- Word2Vec
- GloVe
- FastText

3 Language Models

10. What is a Language Model?

A language model predicts the probability of the next word given previous words.

11. N-gram Models

An n-gram model predicts the next word using the previous $n-1$ words.

12. Limitations of N-gram Models

- Data sparsity
- Large memory requirements
- Poor long-range dependency handling

4 Neural Networks for NLP

13. Why RNNs in NLP?

Recurrent Neural Networks process sequential data and maintain contextual memory.

14. Problems with RNNs

- Vanishing gradients
- Difficulty capturing long-term dependencies

15. LSTM and GRU

LSTM and GRU use gating mechanisms to preserve long-term information and solve vanishing gradient issues.

5 Transformers and LLMs

16. What is a Transformer?

A Transformer is a neural network architecture based on self-attention that processes sequences in parallel.

17. Self-Attention

Self-attention allows each word to focus on other relevant words in a sentence to understand context.

18. Advantages of Transformers

- Parallel computation
- Better long-range dependency modeling
- Faster training

19. What is an LLM?

A Large Language Model (LLM) is a Transformer-based model trained on massive text data to perform multiple NLP tasks.

20. Examples of LLMs

- GPT
- BERT
- LLaMA

21. BERT vs GPT

- **BERT**: Bidirectional, designed for understanding tasks.
- **GPT**: Autoregressive, designed for text generation.

22. Fine-Tuning

Fine-tuning adapts a pretrained model to a specific task using labeled data.

23. Prompt Engineering

Prompt engineering is the practice of designing inputs to guide model outputs without retraining.

6 Training and Evaluation

24. Loss Function in LLMs

Cross-entropy loss is commonly used in language modeling.

25. Perplexity

Perplexity measures how well a language model predicts text; lower values indicate better performance.

26. NLP Evaluation Metrics

- Accuracy
- Precision, Recall, F1-score
- BLEU (machine translation)
- ROUGE (summarization)

7 Advanced Concepts

27. Out-of-Vocabulary Handling

Subword tokenization techniques such as BPE and WordPiece are used.

28. Hallucination in LLMs

Hallucination occurs when a model generates incorrect but confident responses.

29. Reducing Hallucination

- Better prompts
- Fine-tuning
- Retrieval-Augmented Generation (RAG)

30. Retrieval-Augmented Generation (RAG)

RAG combines external document retrieval with text generation to improve factual accuracy.