1. What are Corpora?
2. What are Tokens?
3. What are Unigrams, Bigrams, Trigrams?
4. How to generate n-grams from text?
5. Explain Lemmatization
6. Explain Stemming
7. Explain Part-of-speech (POS) tagging
8. Explain Chunking or shallow parsing
9. Explain Noun Phrase (NP) chunking
10. Explain Named Entity Recognition