Useful NLP Libraries and Networks

Assignment Questions

PW SKILLS

Useful NLP Libraries and Networks

1. What is NLTK?
2. What is SpaCy and how does it differ from NLTK?
3. What is the purpose of TextBlob in NLP?
4. What is Stanford NLP?
5. Explain what Recurrent Neural Networks (RNN) are?
6. What is the main advantage of using LSTM over RNN?
7. What are Bi-directional LSTMs, and how do they differ from standard LSTMs?
8. What is the purpose of a Stacked LSTM?
9. How does a GRU (Gated Recurrent Unit) differ from an LSTM?
10. What are the key features of NLTK's tokenization process?
11. How do you perform named entity recognition (NER) using SpaCy?
12. What is Word2Vec and how does it represent words?
13. Explain the difference between Bag of Words (BoW) and Word2Vec.
14. How does TextBlob handle sentiment analysis?
15. How would you implement text preprocessing using NLTK?
16. How do you train a custom NER model using SpaCy?
17. What is the role of the attention mechanism in LSTMs and GRUs?
18. What is the difference between tokenization and lemmatization in NLP?
19. How do you perform text normalization in NLP?
20. What is the purpose of frequency distribution in NLP?
21. What are co-occurrence vectors in NLP?
22. How is Word2Vec used to find the relationship between words?
23. How does a Bi-LSTM improve NLP tasks compared to a regular LSTM?
24. What is the difference between a GRU and an LSTM in terms of gate structures?
25. How does Stanford NLP's dependency parsing work?
26. How does tokenization affect downstream NLP tasks?
27. What are some common applications of NLP?
28. What are stopwords and why are they removed in NLP?
29. How can you implement word embeddings using Word2Vec in Python?
30. How does SpaCy handle lemmatization?
31. What is the significance of RNNs in NLP tasks?
32. How does word embedding improve the performance of NLP models?
33. How does a Stacked LSTM differ from a single LSTM?
34. What are the key differences between RNN, LSTM, and GRU?
35. Why is the attention mechanism important in sequence-to-sequence models?

Practical

1. How do you perform word tokenization using NLTK and plot a word frequency distribution?
2. How do you use SpaCy for dependency parsing of a sentence?
3. How do you use TextBlob for performing text classification based on polarity?
4. How do you extract named entities from a text using SpaCy?
5. How can you calculate TF-IDF scores for a given text using Scikit-learn?
6. How do you create a custom text classifier using NLTK's Naive Bayes classifier?
7. How do you use a pre-trained model from Hugging Face for text classification?
8. How do you perform text summarization using Hugging Face transformers?
9. How can you create a simple RNN for text classification using Keras?
10. How do you train a Bidirectional LSTM for text classification?
11. How do you implement GRU (Gated Recurrent Unit) for text classification?
12. How do you implement a text generation model using LSTM with Keras?
13. How do you implement a simple Bi-directional GRU for sequence labeling?