

Machine Learning with Python-Practice Exercise

MLA- AIML Batch



NLP Practice Assignments

Day 1

1. What is the purpose of text preprocessing in NLP, and why is it essential before analysis?
2. Describe tokenization in NLP and explain its significance in text processing.
3. What are the differences between stemming and lemmatization in NLP? When would you choose one over the other?
4. Explain the concept of stop words and their role in text preprocessing. How do they impact NLP tasks?
5. How does the process of removing punctuation contribute to text preprocessing in NLP? What are its benefits?
6. Discuss the importance of lowercase conversion in text preprocessing. Why is it a common step in NLP tasks?
7. Explain the term "vectorization" concerning text data. How does techniques like CountVectorizer contribute to text preprocessing in NLP?
8. Describe the concept of normalization in NLP. Provide examples of normalization techniques used in text preprocessing.

Note: Consider the text. It may be a file or prompted inputs.

Python code is mandate for possible Questions.