

Natural Language Processing Fundamentals

Introduction to NLP

Natural Language Processing (NLP) enables machines to understand, interpret, and generate human language.

Tokenization

Tokenization breaks text into words, subwords, or characters for model processing.

Text Cleaning

Includes lowercasing, stopword removal, stemming, and lemmatization.

Language Models

Statistical and neural models that predict the probability of word sequences.

NLP in RAG

NLP techniques help preprocess documents and interpret user queries.