Lesson 5.05
NLP PART 1

Natural Language Processing (NLP)

- Machine's ability to understand, analyse, manipulate, and potentially generate human language

- Real-Life Use Cases
  - Spam Filter in Email Account
  - Auto-Complete in Google Search
  - Auto-Correct in Word Document

NLP Applications

- Sentiment Analysis (Positive / Negative / Neutral)

- Topic Modeling (What are the content pillars in your data?)

- Text Classification (What topic does a text instance belong to?)

Natural Language Toolkit (NLTK)

- Suite of open-source tools created to make NLP in Python easier

NLP Glossary

Term	Definition
NLP (Natural Language Processing)	Machine's ability to understand, analyze, manipulate, and potentially generate human language
NLTK (Natural Language Toolkit)	Most utilized package for handling natural language processing tasks in Python; open-source tools created for the purpose of making NLP process in Python easier
Stemming	The process of reducing inflection or derived words to their word stem or root; aggressively removing the ends of a word to leave only the base
Lemmatizing	The process of grouping together the inflected forms of a word so they can be analyzed as a single term
Stop Words	Commonly used words like "but," "if," "the," etc.
Tokenizing	The process of splitting some string or sentence into a list of words
Vectorizing	The process of encoding text as integers to create feature vectors

Stop Words

Sample text with Stop Words	Without Stop Words
GeeksforGeeks – A Computer Science Portal for Geeks	GeeksforGeeks , Computer Science, Portal ,Geeks
Can listening be exhausting?	Listening, Exhausting
I like reading, so I read	Like, Reading, read

Stemming Versus Lemmatization