Natural Language Processing Fundamentals

Course Summary

Description

You'll be introduced to natural language processing and its applications through examples and exercises. This will be followed by an introduction to the initial stages of solving a problem, which includes problem definition, getting text data, and preparing it for modeling. With exposure to concepts like advanced natural language processing algorithms and visualization techniques, you'll learn how to create applications that can extract information from unstructured data and present it as impactful visuals. Although you will continue to learn NLP-based techniques, the focus will gradually shift to developing useful applications. In these sections, you'll understand how to apply NLP techniques to answer questions as can be used in chatbots.

By the end of this course, you'll be able to accomplish a varied range of assignments ranging from identifying the most suitable type of NLP task for solving a problem to using a tool like Spacy or Gensim for performing sentiment analysis. The course will easily equip you with the knowledge you need to build applications that interpret human language.

Objectives

By the end of this course, students will be able to:

- Obtain, verify, and clean data before transforming it into a correct format for use
- Perform data analysis and machine learning tasks using Python
- Understand the basics of computational linguistics
- Build models for general natural language processing tasks
- Evaluate the performance of a model with the right metrics
- Visualize, quantify, and perform exploratory analysis from any text data

Topics

- Introduction to NLP Text
- Extraction Methods from unstructured text
- Building A Simple Classifier
- Collecting Text Data
- Topic Modeling
- Text Summarization and Text Generation
- Vector Representation
- Sentiment Analysis

Audience

Natural Language Processing Fundamentals is designed for novice and mid-level data scientists and machine learning developers who want to gather and analyze text data to build an NLP-powered product. It'll help you to have prior experience of coding in Python using data types, writing functions, and importing libraries. Some experience with linguistics and probability is useful but not necessary.

Duration

Three Days
Course Outline

I. Introduction to NLP
   A. Text Analytics and NLP
   B. Various Steps in NLP
   C. Kick Starting an NLP Project

II. Extraction Methods from unstructured text
   A. Types of Data
   B. Cleaning Text Data
   C. Feature Extraction from Texts
   D. Feature Engineering

III. Building A Simple Classifier
   A. Machine Learning
   B. Developing a Text Classifier
   C. Building Pipelines for NLP Projects
   D. Saving and Loading Models

IV. Collecting Text Data
   A. Collecting Data by Scraping Web Pages
   B. Requesting Content from Web Pages
   C. Dealing with Semi-Structured Data

V. Topic Modeling
   A. Topic Discovery
   B. Topic Modeling Algorithms

VI. Text Summarization and Text Generation
   A. What is Automated Text Summarization?
   B. High-Level View of Text Summarization
   C. TextRank
   D. Summarizing Text Using Gensim
   E. Summarizing Text Using Word Frequency
   F. Generating Text with Markov Chains

VII. Vector Representation
   A. Vector Definition
   B. Why Vector Representations?

VIII. Sentiment Analysis
   A. Why is Sentiment Analysis Required?
   B. Growth of Sentiment Analysis
   C. Tools Used for Sentiment Analysis
   D. TextBlob
   E. Understanding Data for Sentiment Analysis
   F. Training Sentiment Models