

Generative AI for Natural Language Processing

Foundational Data Course



Course Timeframe

4-week part-time course. 2 lectures per week.

Course Delivery

Live online.
Office hours are included for course duration.

Who is This Course For?

- Individuals with a strong educational background and professionals outside of STEM with little to no experience in data who have a strong interest in learning how to leverage data science techniques and skills
- Junior business analysts, data analysts, market intelligence analysts, product operations, design operations, project managers, and professionals looking to get a foundational understanding of AI/data science techniques and learn how to leverage AI tools to solve simple to intermediate real-world business problems

Who is This Course Not For?

Experienced data science, data analytics, or data engineering professionals

Prerequisites

Completion of the "Data Wrangling with Python" course or equivalent programming experience

Course Learning Objectives

Learn how to use generative AI and LLM (Large Language Models) offered by OpenAI to process and extract data from natural language text and enhance your data science models with deeper insights. Practice best practices for prompting the models and gain experience in loading the results into Python for further analysis. Augment your existing numerical analyses by incorporating data derived from textual data.

By The End of This Course, Students Will...

- Describe the capabilities and limitations of large language models
- Write effective prompts for LLMs
- Use the OpenAI API to query LLMs
- Process natural language with LLMs
- Extract sentiment and topics from natural language text
- Ingest the results of LLMs into Python for further analysis

Use Case Examples

- Separate positive and negative product reviews, based only on their text
- Analyze customer survey responses, identify key topics, and summarize feedback
- Automate document processing: summarize, organize, and refine content, with a query capability
- Automatically and iteratively refine the analysis to ensure desired results
- Conduct sentiment analysis of ongoing conversational data stream from chat boxes
- Aggregate customer experiences into consistent categories, based on written descriptions

[Click here to learn more](#)



Generative AI for NLP Course Syllabus

Foundational Data Course



An Eight-Module Structured Learning Path

Module 1: Using Web APIs

Using the requests package, parsing JSON, authenticated APIs, caching and limiting requests

Module 2: Large Language Models

Building a next-word predictor, generating whole documents, using transformers and GPT models

Module 3: Using the GPT Models

Accessing the API, the OpenAI Python package, handling errors, basic prompt engineering

Module 4: Using LLMs for Data Analysis

Extracting structured information, ensuring the output format, dealing with missing data

Module 5: Framing Natural Language Problems: Sentiment Analysis

Comparing traditional NLP to LLMs, definition and application of sentiment analysis, automating the analysis

Module 6: Topic Modeling with LLMs

Summarizing documents, generating topics, categorizing texts

Module 7: Vector Representations and Databases

Text embeddings and semantic vectors, measuring document similarity, vector databases, [retrieval-augmented generation (to be developed)]

Module 8: Other Uses of LLMs

Summarizing, generating, and translating documents; LLMs as programming assistants, function calling

Includes hands on exercises and mini project