Large Language Models, Artificial Intelligence and Data Science

Syllabus: CS211 Section 301 Spring 2025

Info

3 Credits
Tue 2:10pm - 5:00pm
Room C305

Instructor Information

calvin_williamson@fitnyc.edu
office: B831 Science and Math
office hours: M 1-3, T 11-1, W 11-12

Description

This course provides an introduction to large language models and their capabilities in artificial intelligence and data science. Through theory and hands-on labs, students will gain literacy in model architecture, training techniques, applications in programming, content creation, and more. No prior programming experience is required to take this course. Prerequisite(s): Math Proficiency

Outcomes

1. Explain the evolution, capabilities, and limitations of large language models.
2. Analyze the core components of LLMs including model architecture, parameters, and training techniques.
3. Assess strategies like prompt engineering and fine-tuning to optimize LLM performance.
4. Explore mathematical foundations of vector databases and storage including semantic similarity
5. Apply LLMs to natural language tasks like classification, summarization, and question answering.
6. Utilize LLMs to assist with programming, data science, content creation, and other domains.
7. Examine ethical implications of LLMs including bias, misinformation, and legal issues.

Course Materials

We will be using Google Colab, OpenAI ChatGPT, Anthropic Claude, Google Gemini, and other LLM tools for all work in this course. Since these are web-based there is NO OTHER SOFTWARE required for the course besides a web browser.

Topics

Introduction to Large Language Models (LLMs)

- LLM Examples (ChatGPT, Gemini, Claude, DeepSeek)
- Applications: NotebookLM
- Multimodal Model Examples (Gemini 2.5)

Introduction to Python for Artificial Intelligence

- Google Colab Notebook
- Using LLM as Coding Assistant
- Calculations
- Variables
- DataTypes
- Lists
- Dictionaries
- Functions
- Dataframes
- f-Strings

Prompt Engineering Techniques

- Completions, APIs
- System Prompts
- Context Size
- Prompt Structure
- Few-shot and Zero-shot Learning
- Structured Outputs
- Prompt Chaining
- Roles and Personas
- Chain of Thought
- Reasoning Models
- Chat Conversations, Memory
- Multimodal Prompt Techniques

LLM Applications

- Classification
- Artifacts
- Sentiment Analysis
- Question Answering
- Text Summarization
- Named Entity Recognition (NER)

Image Generation Models

- Prompting for Image Generation
- FLUX model

Tools, Agents

- Tools and APIs
- Model Context Protocol (MCPS)

Agentic Software Tools

- Artifacts
- Visual Studio Code (Cline)
- CLI(Command Line) Tools

LLM and Data Science

- Databases And AI-Assisted SQL
- Data Science Agents

Evaluation

Your grade will come from these parts:

- Quizzes (85%)
- In Class/Homework Credits (15%)

Each of these parts is described in more detail below

Quizzes

Your quiz grade will come from 5 quizzes roughly covering 2 or 3 weeks material eachThis quizzes are 30-45 minutes each and are usually 5 or 6 questions each.These quizzes are with no notes, no internet, no phone, no software, no AI tools.Pen and paper and calculator only. They are some multiple choice, some short answer, some true false.

In Class/Homework (1 or 2 per class)

These are credits you obtain for demonstrating you have completed assigned problems. Some of these will come from homework assignments that you show me at the beginning of the class, some of these will come from in class assignments that are done during class and you show as you complete them. You will earn a credit for each successful assignment completion. You must be in attendance to earn these problem credits.

There is NO FINAL EXAM.

AI Policy

All uses of chatbots are encouraged, and there is no restriction on their use. This is especially for topics about large language models (ChatGPT, Gemini, Claude, etc).