Northeastern College of Engineering

CSYE7380 Theory & Practice of AI Generative Models

Course Information
Course Title: Theory & Practice of AI Generative Models
Course Number: CSYE7380
Term and Year:
Spring 2025 Credit
Hour: 4
Course Format: On-Ground

Instructor Information
Full Name: William Claster
Email Address: w.claster@northeastern.edu

Course Prerequisites
INFO 6205 with a minimum grade of B

Course Description

This course provides a comprehensive exploration of AI generative models, with a focus on Large Language Models (LLMs) and their applications. Students will gain hands-on experience with state-of-the-art techniques in natural language processing, deep learning, and agentic AI systems. We will learn various methods for aligning and fine-tuning LLMs to support business and organizational aims. The course emphasizes both theoretical foundations and production-ready implementations, with substantial focus on LangChain and LangGraph for building autonomous AI agents. Students will complete a final project demonstrating practical deployment of these technologies.

Prerequisites: Basic knowledge of machine learning and Python programming.

Course Structure: 14 weeks of lectures with 4 programming assignments, 1 midterm exam, and 1 final project.

Course Schedule/Topics Covered

Week 1: Foundations & LangChain Introduction

Agentic AI (1.5 hours) • Introduction to LangChain architecture • Basic agents: Zero-shot ReAct • Tools and tool calling • Lab: Build first agent with calculator and search tools

Theory (1.5 hours) • Neural network fundamentals review • Backpropagation and gradient descent • Introduction to PyTorch • Tokenization and embeddings

Week 2: LangGraph & Language Modeling

Agentic AI (1.5 hours) • LangGraph state machines • Conditional edges and cycles • Checkpointing and persistence • Lab: Multi-step workflow agent

Theory (1.5 hours) • Language modeling fundamentals • Autoregressive generation • Perplexity and evaluation metrics • Word2Vec to contextual embeddings

Week 3: Multi-Agent Systems & Transformers

Agentic AI (1.5 hours) • Multi-agent orchestration patterns • Supervisor and delegation patterns • Agent communication protocols • Lab: Build supervisor-worker system

Theory (1.5 hours) • Transformer architecture deep dive • Attention mechanisms (self, cross, multi-head) • Positional encoding • BERT vs GPT architectures

Week 4: LLMs & Agent Memory Systems

Agentic AI (1 hour) • Memory patterns for agents • Short-term vs long-term memory • Vector memory integration

Theory (2 hours) • GPT-3/4 architecture and scale • Chain-of-thought reasoning • In-context learning • Few-shot vs zero-shot learning

Week 5: Fine-tuning & Agent Planning

Agentic AI (1 hour) • Plan-and-execute agents • Tree of Thoughts implementation • Reflexion pattern

Theory (2 hours) • Fine-tuning fundamentals • Catastrophic forgetting • Dataset preparation • Evaluation strategies • Weights and Biases for tracking

Week 6: VAE & Structured Generation

Agentic AI (1 hour) • Structured outputs for agents • JSON mode and function calling • Grammar-guided generation

Theory (2 hours) • Variational Autoencoders • Latent space representations • VAE loss functions • Applications in generation

Week 7: Advanced Techniques & Security

Agentic AI (1 hour) • Agent error handling • Fallback chains • Human-in-the-loop patterns

Theory (2 hours) • LoRA and QLoRA • Prompt injection attacks and defense • Jailbreaking and red-teaming • Lab: Build injection detection

Week 8: RAG Systems Deep Dive

Agentic AI (1.5 hours) • RAG + Agent integration • Self-RAG patterns • Agentic retrieval strategies

Theory (1.5 hours) • Embeddings and vector databases • Hybrid search strategies • Reranking algorithms • Midterm review session

Week 9: MIDTERM EXAM & Production Basics

Midterm Exam (1.5 hours) • Closed book, paper and pencil only • One formula sheet allowed (8.5x11, both sides) • Covers material from Weeks 1-8

Production Deployment Introduction (1.5 hours) • Model quantization basics • API design overview • Container deployment intro

Week 10: RLHF & Alignment Basics

Agentic AI (1 hour) • Agents with reward models • Self-critique agents • Constitutional AI agents

Theory (2 hours) • RLHF fundamentals • Reward model training • PPO for language models • Human preference collection

Week 11: Modern Alignment Methods

Agentic AI (1 hour) • RLAIF implementation • Agent-based evaluation

Theory (2 hours) • Direct Preference Optimization (DPO) • Constitutional AI deep dive • Synthetic preference generation • Lab: DPO fine-tuning

Week 12: Advanced Architectures & Cost Optimization

Theory (1.5 hours) • Mixture of Experts (MoE) • Long context models • Flash Attention

Production Skills (1.5 hours) • Cost optimization at scale • Model selection strategies • Caching and batching

Week 13: LLM Testing & Evaluation

Core Content (2 hours) • Custom benchmark creation • Regression testing suites • Data contamination detection • IP and PII testing

Project Presentations (1 hour) • Student project progress presentations

Week 14: Multimodal Models & Course Wrap-up

Core Content (1.5 hours) • Vision-language models • Speech integration • Current research frontiers

Final Project Presentations (1.5 hours) • Final project demonstrations

Major Assignments

Each assignment builds toward developing and improving production-ready AI systems:

Assignment 1 (Due Week 3): Basic LangChain Agent

• Build an agent with custom tools

• Implement memory management

• Include error handling

Assignment 2 (Due Week 6): Fine-tuned Model with Evaluation

• Fine-tune a model for specific task

• Create evaluation metrics

• Compare with baseline

Assignment 3 (Due Week 8): Secure RAG System with Agent

• Implement RAG pipeline

• Add prompt injection defense

• Include agent orchestration

Assignment 4 (Due Week 11): Production Deployment with Monitoring

• Deploy model via API

• Implement cost tracking

• Add performance monitoring

Final Project (Due Week 14): Enterprise-Ready LLM Application

• Choose from: RAG system for documents, Customer support agent, Financial document analyzer, or Multi-agent research system

• Must include evaluation metrics, cost analysis, and deployment strategy

Midterm Exam Information

Week 9 - In-class exam (1.5 hours) • Closed book, paper and pencil only • One formula sheet allowed (8.5x11", both sides) • No phones, calculators, or electronic devices

Exam Coverage: • Mathematical foundations (attention calculations, loss functions, embeddings) • Conceptual understanding (architectures, fine-tuning, agent design) • Practical applications (security, RAG systems, memory patterns)

Grading Scale

Grade	Percentage	Grade	Percentage	Grade	Percentage
A	95-100%	B+	87-89.9%	C+	77-79.9%
A-	90-94.9%	B	84-86.9%	C	74-76.9%
		B-	80-83.9%	C-	70-73.9%
				F	69.9% or below

Grade Distribution

• Midterm Exam: 15%

• Assignments: 30% (4 assignments × 7.5% each)

• Final Project: 30%

• Class Participation: 10%

• Paper Presentations: 5%

• Quizzes: 10%

Required Readings

Key papers will be assigned throughout the course, including: • "Attention Is All You Need" (Vaswani et al., 2017) • "Language Models are Few-Shot Learners" (Brown et al., 2020) • "Direct Preference Optimization" (Rafailov et al., 2023) • "ReAct: Synergizing Reasoning and Acting" (Yao et al., 2023) • Additional current papers as they become available

Attendance Policy

Students registered in MGEN courses (INFO, CSYE, and DAMG) are allowed **a maximum of 2 absences per course, with 3 or more absences resulting in an automatic 'F' for that course.** Students are expected to inform their instructors of any absences in advance of the class; if a student is sick long-term or experiences a medical issue that prevents class attendance, it is strongly encouraged that they speak with their Academic Advisor (coe-mgen-gradadvising@northeastern.edu) to learn more about the Medical Leave of Absence. Should a student anticipate being unable to attend 3 or more classes, they should discuss their situation with their Academic Advisor to explore other types of leave in accordance with the University's academic and global entry expectations. International students should review the Office of Global Services webpage to understand their visa compliance requirements.

Teaching Assistants (TAs) or Instructional Assistants (IAs) will be present at each class to collect student attendance.

Late Work Policy

Students must submit assignments by the deadline in the time zone noted in the syllabus. Students must communicate with the faculty prior to the deadline if they anticipate work will be submitted late. Work submitted late without prior communication with faculty will not be graded.

End-of-Course Evaluation Surveys
Your feedback regarding your educational experience in this class is particularly important to the College of Engineering. Your comments will make a difference in the future planning and presentation of our curriculum.

At the end of this course, please take the time to complete the evaluation survey at **https://neu.evaluationkit.com**. Your survey responses are **completely anonymous and confidential**. For courses 6 weeks in length or shorter, surveys will be open one week prior to the end of the courses; for courses greater than 6 weeks in length, surveys will be open for two weeks. An email will be sent to your Northeastern University Mail account notifying you when surveys are available.

Academic Integrity
A commitment to the principles of academic integrity is essential to the mission of Northeastern University. The promotion of independent and original scholarship ensures that students derive the most from their educational experience and their pursuit of knowledge. Academic dishonesty violates the most fundamental values of an intellectual community and undermines the achievements of the entire University. As members of the academic community, students must become familiar with their rights and responsibilities. In each course, they are responsible for knowing the requirements and restrictions regarding research and writing, examinations of whatever kind, collaborative work, the use of study aids, the appropriateness of assistance, and other issues. Students are responsible for learning the conventions of documentation and acknowledgment of sources in their fields. Northeastern University expects students to complete all examinations, tests, papers, creative projects, and assignments of any kind according to the highest ethical standards, as set forth either explicitly or implicitly in this Code or by the direction of instructors.

Go to http://www.northeastern.edu/osccr/academic-integrity-policy/ to access the full academic

integrity policy.

MGEN Student Feedback
Students who would like to provide the MGEN unit with anonymous feedback on this particular course, Teaching Assistants, Instructional Assistants, professors, or to provide general feedback regarding their program, may do so using this survey:
https://neu.co1.qualtrics.com/jfe/form/SV_cTIAbH7ZRaaw0Ki

University Health and Counseling Services
As a student enrolled in this course, you are fully responsible for assignments, work, and course materials as outlined in this syllabus and in the classroom. Over the course of the semester if you experience any health issues, please contact UHCS.

For more information, visit https://www.northeastern.edu/uhcs.

Student Accommodations
Northeastern University and the Disability Resource Center (DRC) are committed to providing disability services that enable students who qualify under Section 504 of the Rehabilitation Act and the Americans with Disabilities Act Amendments Act (ADAAA) to participate fully in the activities of the university. To receive accommodations through the DRC, students must provide appropriate documentation that demonstrates a current substantially limiting disability.

For more information, visit https://drc.sites.northeastern.edu.

Library Services
The Northeastern University Library is at the hub of campus intellectual life. Resources include over 900,000 print volumes, 206,500 e-books, and 70,225 electronic journals.

For more information and for education specific resources, visit
https://library.northeastern.edu Network Campus Library Services: Northeastern University Library Global Campus Portals

24/7 Canvas Technical Help
For immediate technical support for Canvas, call 617-373-4357 or email

help@northeastern.edu Canvas Student Resources:

https://canvas.northeastern.edu/student-resources/For assistance with my Northeastern

e-mail, and basic technical support:

Visit ITS at
https://its.northeastern.edu
Email: help@northeastern.edu
ITS Customer Service Desk: 617-373-4357

Diversity and Inclusion
Northeastern University is committed to equal opportunity, affirmative action, diversity, and social justice while building a climate of inclusion on and beyond campus. In the classroom, members of the University community work to cultivate an inclusive environment that

denounces discrimination through innovation, collaboration, and an awareness of global perspectives on social justice.

Please visit http://www.northeastern.edu/oidi/ for complete information on Diversity and Inclusion

Title IX

Title IX of the Education Amendments of 1972 protects individuals from sex or gender-based discrimination, including discrimination based on gender-identity, in educational programs and activities that receive federal financial assistance.

Northeastern's Title IX Policy prohibits Prohibited Offenses, which are defined as sexual harassment, sexual assault, relationship or domestic violence, and stalking. The Title IX Policy applies to the entire community, including male, female, transgender students, faculty, and staff.

In case of an emergency, please call 911.

*Please visit **https://www.northeastern.edu/ouec** for a complete list of reporting options and resources both on- and off-campus.*