**Boston University** Metropolitan College

# Generative AI

MET CS 788

Course Format (On Campus)

Reza Rawassizadeh
rezar@bu.edu
Office hours: by appointment

**Course Description**

This course focuses on technical details and the mathematical foundations behind generative AI models. In this course, first, we learn statistical concepts required for generative artificial intelligence. Next, we review regressions and optimization methods. Afterward, we review traditional neural network architectures, including perceptron and multilayer perceptron. Next, we move to Convolutional Neural Networks and Recurrent Neural Networks and close this part with Attention and Transformers.

The second part of the course focus on generative neural networks. We start with traditional self-supervised learning algorithms (Self Organized Map and Restricted Boltzmann Machine). Then, we explore Auto Encoder architectures and Generative Adversarial Networks. Afterward, we moved toward architectures that construct generative models, including recent advances in NLP e.g., BERT, InstructGPT (ChatGPT architecture). Finally, we describe Neural Radiance Field, 3D Gaussian Splatting, and text-2-image models.

**Books**

Some chapters of the instructor's ongoing book: https://github.com/Rezar/MLBook
Besides, students who are willing to use extra resources can check the following books:

- An Introduction to Statistical Learning: with Applications in R (James et al.)
- Dive into Deep Learning (Zhang et al.)

**Course Requirements**

- Students should be familiar with Python programming language and some mathematics required for machine learning.
- Students must CS 677 and it is highly recommended to be familiar with neural networks.
- This course includes lots of theories, and except for assignments, we do not go into implementation details. Students should be able to learn Keras, Tensor Flow, or Pytorch on their own. Besides, they need to be able to use cloud services such as Google Colab environment to build and train their neural network model.

**Class Policies**

1) **Attendance & Absences** – Class attendance is not mandatory but highly recommended. All quizzes and the final exam will be done inside the class, and taking online exams are impossible.
2) **Assignment Completion & Late Work** – 40% - 50% of the final grade comes from assignment and project delivery. Late submission of homework is associated with a penalty of 10% grade reduction for any single day.
3) **Quiz and Final exam –** 30%- 40% of the final grade comes from quizzes and the final exam. Quiz and final exams focused on concepts and not coding.
4) **Final Project –** 10% -20% of the final grade comes from final project
5) **Academic Conduct Code** – "Cheating and plagiarism will not be tolerated in any Metropolitan College course. They will result in no credit for the assignment or examination and may lead to disciplinary actions. Please take the time to review the Student Academic Conduct Code: http://www.bu.edu/met/metropolitan_college_people/student/resources/conduct/code.html.

**Grading Criteria**

Student grades are the sum of their assignments, final exam, and quizzes. The resulting grade will be calculated as follows.

| | |
|---|---|
| A | 95-100 |
| A- | 90-94.99 |
| B+ | 85-89.99 |
| B | 80-84.99 |
| B- | 75-79.99 |
| C+ | 70-74.99 |
| C | 65-69.99 |
| F | <65 |

**Class Meetings, Lectures & Assignments**

*Lectures, Readings, and Assignments are subject to change and will be announced in class as applicable within a reasonable time frame.*

| Date | Topic | Assignments Due |
|---|---|---|

| | | |
|---|---|---|
| Session 1 | Introduction and Statistics I (Statistical Distribution Concepts and Requirements for Generative AI) | NA |
| Session 2 | Statistics II (Statistical Distribution Comparison Methods and Challenges | 14 days after Session 2 |
| Session 3 | Characteristics of Linear versus Non-linear Models. | NA |
| Session 4 | Mathematical Concepts for Optimization (Gradient, Jacobian, Hessian, and Taylor Series). | 14 days after Session 4 |
| Session 5 | Review the basics of Neural Networks (MLP, Activation Functions, Backpropagation, Cost Functions) | NA |
| Session 6 | Review of Convolutional Neural Networks and Recurrent Neural Networks | NA |
| Session 7 | Review of Seq2Seq, Attention Mechanisms, Positional Encoding, Transformer Architecture. | 7 days after Session 7 |
| Session 8 | Representation Learning Concepts, Self-Organized Maps, Boltzmann Machine, Restricted Boltzmann Machine | 14 days after Session 8 |
| Session 9 | Autoencoders, Generative Adversarial Networks, GAN challenges, GAN Models | NA |
| Session 10 | Large Language Models I (BERT models, GPT models, Post ChatGPT models) | NA |
| Session 11 | Large Language Models II (Prompt engineering methods, NLP evaluations, LLM, RAG) | 14 days after Session 10 |
| Session 12 | Text-to-Image Models | NA |
| Session 13 | Neural Radiance Field, 3D Gaussian Splatting | NA |
| Session 14 | Neural Network Regularization and Compression Approaches (Quantization, Compression, Pruning, Low-Rank Adaptation) | NA |