INFO 6105 & Data Science Engineering Methods
Fall 2024

Course Information
Course Title: Data Science Engineering Methods
Course Number: INFO 6105
Term and Year: Fall 2024
Credit Hour: 4
CRN: 20823
Course Format: Traditional

Instructor Information
Full Name: Ahmed Abdeen Hamed
Email Address: a.hamed@northeastern.edu
Office Hours: M/W 130PM-3PM

Instructor Biography
Professor of Data Science, LLMs, and Artificial Intelligence

Teaching Assistant Information
Full Name: N/A
Email Address:N/A
Office Hours:N/A

Course Prerequisites
N/A

Course Description
This course covers the essential techniques in machine learning and data science engineering. It explores a wide range of machine learning algorithms, with practical examples of their implementation, evaluation, and best practices. The course builds a foundational understanding of how learning models are developed from complex data pipelines, both in theory and practice. Key topics include supervised learning (covering parametric and nonparametric algorithms, support vector machines, kernel methods, neural networks, and deep learning) and unsupervised learning (including clustering, dimensionality reduction, and recommender systems). The curriculum is enriched with numerous real-world case studies.
Proficiency in Python is required as part of the grading. Students must demonstrate their ability to prepare data for learning, and perform training, testing, and evaluation using either Python, though all examples

and solutions in assignments will be provided in Python. Assignments will include exercises to reinforce conceptual understanding, along with exams, and a required term project.

Course Learning Outcomes

Understand Fundamental Machine Learning Concepts: Students will be able to explain the core principles and techniques of machine learning and data science engineering, including both supervised and unsupervised learning algorithms.

Implement Machine Learning Algorithms: Students will gain hands-on experience in implementing a variety of machine learning algorithms using Python, demonstrating an ability to apply these techniques to real-world datasets.

Evaluate and Compare Learning Models: Students will learn to assess the performance of different machine learning models through appropriate evaluation metrics and techniques, and to compare models to determine the most suitable for specific tasks.

Develop Complex Data Pipelines: Students will be able to construct and manage complex data pipelines, integrating various preprocessing, training, and testing stages necessary for machine learning applications.

Apply Learning Models to Real-World Problems: Through case studies and practical assignments, students will apply machine learning models to solve real-world problems, gaining insight into the challenges and considerations involved.

Proficiency in Python for Data Science: Students will demonstrate proficiency in using Python for data manipulation, model training, testing, and evaluation, as well as in documenting their work through a portfolio blog.

Develop and Present a Term Project: Students will design, implement, and present a comprehensive term project that showcases their ability to apply machine learning techniques to a complex problem, from data preparation to model evaluation.

Communicate Technical Concepts: Students will enhance their ability to communicate machine learning and data science concepts effectively through written assignments, manuscript writing for a conference publication, and project presentations.

Required Tools and Course Textbooks.

Textbook: Data Mining Concepts and Techniques

Programming Language: Python

Course Schedule/Topics Covered.

Week	Date	In Class Topic	Assignment Due
1	**09/04**	Chapter 1: Introduction to Data Mining and Data Science	
2	**09/11**	Chapter 2: Getting to Know Your Data	
3	**09/18**	Chapter 3: Data Preprocessing Importance and Techniques	Assignment 1
4	**09/25**	Chapter 6: Mining Frequent Patterns, Association and Correlations: Basic Concepts and Methods	
5	**10/2**	Chapter 7: Advanced Frequent Pattern Mining	Assignment 2
6	**10/9**	Chapter 8: Chapter 8. Classification: Basic	

		Concepts	
7	**10/16**	Chapter 9. Classification: Advanced Methods	Assignment 3
8	**10/23**	Chapter 10. Cluster Analysis: Basic Concepts and Methods	
9	**10/30**	Chapter 11: Advanced Cluster Analysis	Assignment 4
10	**11/6**	Chapter 12. Outlier Analysis	
11	**11/13**	Chapter 13: Data Mining Trends and Research Frontiers	Final Project
12	**11/20**	Networks	
13	**11/27**	Network-driven Classification	
14	**12/4**	Network-driven Clusters	
15	**12/11**	Final Project Due	

Rubric and Grading

- Attendance – 10 %
- 4 Assignments: 40% (each is 10%)
- Midterm Eaxm: 20%
- Final Project: Code Complettion, Manuscript, submitted to a workshop 30%

Grading Scale

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

Attendance/Late Work Policy

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_cTlAbH7ZRaaw0Ki

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.*