Northeastern College of Engineering

INFO 6105 INFO Data Science Engineering Methods and Tools
[FALL 2025]

Course Information
Course Title: Data Science Engineering Methods and Tools
Course Number: INFO 6105
Term and Year: Fall 2025
Credit Hour: 4
CRN: 17651
Course Format: On-Ground
Timing: Thursday's 3:15 – 6:35 pm PST.
Lecture location: 1523 (Vancouver Campus)

Instructor Information
Full Name: Rushdi Alsaleh, PhD
Email Address: r.alsaleh@northeastern.edu
Office Hours: On-campus:
          Wed 4:00-5:30 pm
          Thurs 6:30-7:30 pm
          Or
          Online through the link:
          https://calendly.com/ralsaleh/15-min-gn
          on Mon 9:00-10:00 am & 4:00-5:00 pm
          Tues 9:30 am – 12:30 pm
          Fri 10:30 am – 2:00 pm
          Or
          Anytime by appointment

Instructor Biography
Professor of Data Science, and Artificial Intelligence, with more than 12 years of experience in industry and has led multiple high-impact data-driven projects with local and global companies.

Teaching Assistant Information
Full Name: Yi Zhen
Email Address: yi.zhen1@northeastern.edu
Office Hours:TBD

Course Prerequisites
Please consult school listings for prerequisites.

Course Description
This course provides an introduction to machine learning using Python, the open source, programming language extensively adopted by the machine learning community and industry. Machine learning is at the center of a powerful movement. Many industries depend on practitioners of machine learning to create products that parse, reduce, simplify and categorize data, and then extract actionable intelligence from that data. Professionals who are familiar with machine learning, a key technology driving Big Data, secure a competitive edge in exciting careers in the data sciences. In this course, you will use Python to learn machine learning concepts, terms and methodology, and gain an intuitive understanding of the mathematics underlying it by building actual applications. The techniques you'll learn can be a starting point to build real-world applications such as search engines, image analysis, bioinformatics, industrial automation, speech recognition, and more.

This course establishes a basic understanding of supervised learning and Bayesian classifiers using the histogram as a starting point. It then covers the design and application of practically useful classifiers such as k-nearest neighbors, linear machines and decision trees. You will also learn concepts in unsupervised learning and clustering algorithms such as expectation maximization and k-means clustering. The course concludes with the application of neural networks in machine learning.

The course uses examples to guide you through foundational concepts, often employing live algorithms to facilitate visual understanding. Pseudocode will be provided for most of the algorithms covered. You are encouraged to use the pseudocode as a reference to create your own programs in Python. The class has in-class quizzes to gauge learning and group activities including discussion. Homework assignments involving programming in Python are designed for in-depth practice.

Course Learning Outcomes
Learning outcomes common to all College of Engineering Graduate programs:
1.   An ability to identify, formulate, and solve complex engineering problems.
2.   An ability to explain and apply engineering design principles, as appropriate to the program's educational objectives.
3.   An ability to produce solutions that meet specified end-user needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors.

The Information Systems Program accepts students of different engineering backgrounds with minimum programming skills and produces first class Information Systems engineers that operate at the intersection of real-world complexity, software development, and IT management. Graduating students will be able to construct end-to-end advanced software applications that meet business needs.

Specific Learning Outcomes for the Information Systems program:

1. Create a strong technical foundation through diverse, high-level courses
2. Built crucial interpersonal skills needed to succeed in any industry
3. Foster a deep level of applied learning through project-based case studies

Required Tools and Course Textbooks.

The material for this course is designed from sctarch and the following textbooks are recommended:

Recommended Textbooks:

- *Data Mining for Business Analytics: Concepts, Techniques and Applications in Python*. 2019. Galit Shmueli, Peter C. Bruce, Peter Gedeck, Nitin R. Patel. ISBN: 978-1-119-54984-0
- *An introduction to statistical learning: With applications in python*. James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J., Springer Nature.
- *Data Mining: Concepts and Techniques*, 3rd ed. The Morgan Kaufmann Series in Data Management Systems Morgan Kaufmann Publishers, 2011
- *The Elements of Statistical Learning Data Mining, Inference, and Prediction, Second Edition*. Hastie, T., Tibshirani, R., & Friedman, J. 2009, Springer Nature.
- *Machine Learning*, Tom M. Mitchell, McGraw-Hill, 1997, ISBN-10: 0070428077, ISBN-13: 978-0070428072.
- *Pattern Recognition and Machine Learning*, Christopher M. Bishop, Springer, 2007, ISBN-10: 0387310738, ISBN-13: 978- 0387310732.
- *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd Edition, Ian H. Witten, et al., Morgan Kaufmann, 2011, ISBN-10: 0123748569, ISBN-13: 978-0123748560.
- *Data Science Using Python and R*. Chantal D. Larose, Daniel T. Larose, 2019. ISBN: 978-1-119-52681-0. Wiley

Software

- **Python Anaconda**
    https://www.continuum.io/anaconda-overview
- **R (Statisical programming language)**
    https://www.r-project.org/
- **RStudio (IDE)**
    https://posit.co/download/rstudio-desktop/

Course Schedule/Topics Covered.

Week	Date	In Class Topic	Assignment Due
1	09/04	Introduction: Concepts in Machine Learning and Predictive Analytics	Python certificate for Data Science (DS) (Module 1)
2	09/11	Intro to python/R and Data Pre-processing, Data Visualization, EDA	Python Certificate for DS (Module 2)
3	09/18	Linear Models, Model Inference and Interpretation, Regression, Logistic Regression	Python Certificate for DS (Module 3)

			Assignment 1
4	**09/25**	Supervised Learning: Decision Tree, CART, Bagging and Random Forest	Python Certificate for DS (Module 4)
5	**10/2**	Supervised Learning: Decision Tree, CART, Bagging and Random Forest Supervised Learning: Support Vector Machine (SVM)	Python Certificate for DS (Module 5) Assignment 2
6	**10/09**	Supervised Learning: Support Vector Machine (SVM) Case studies	Assignment 3 Python Certificate for DS Submission
7	**10/16**	Midterm Exam	Midterm Exam Python Certificate for Data Visualization (DV) (module 1)
8	**10/23**	Research progress report Neural Netwrks theories and implementation	Python Certificate for DV (module 2)
9	**10/30**	Neural Netwrks theories and implementation Unsupervised learning: finding Groups of Data (Clustering)	Assignment 4 Python Certificate for DV (module 3)
10	**11/6**	Unsupervised learning: finding Groups of Data (Clustering)	Python Certificate for DV (module 4)
11	**11/13**	Evaluating and Improving Model Performance Overfitting and Regularization Applying Machine Learning Guidance and Practical Issues	Assignment 5 Python Certificate for DV (module 5)
12	**11/20**	Classification: Nearest Neighbors, Naïve Bayes Applying Machine Learning Guidance and Practical Issues Into to deep learning	Python Certificate for DV Submission
13	**11/27**	Fll Break – No class	
14	**12/4**	Final Term Project Presentation	Project Presentation
15	**12/11**	Final Exam	Final Exam
		Final Grade submission Deadline (Dec. 16th)	

Assignment Grading
Please insert all assignment grades and weights for the course. Example below:
- Attendance and participation – 5%
- Pyhton Certificates – 10%
- Assignments – 25%
- Midterm Exam – 20%

- Term Project – 20%
- Final Exam– 20%

Grading Scale

Percentage Range	Letter Grade	Grade Point Equivalent
95.0–100.0%	A	4.000
90.0–94.9%	A-	3.667
87.0–89.9%	B+	3.333
84.0–86.9%	B	3.000
80.0–83.9%	B-	2.667
77.0–79.9%	C+	2.333
74.0–76.9%	C	2.000
70.0–73.9%	C-	1.667
69.9% and Below	F	0.000

Incomplete Grades

An incomplete grade may be reported by the instructor when a student has failed to complete a major component of a required course, such as homework, a quiz or final examination, a term paper, or a laboratory project. Students may make up an incomplete grade by satisfying the requirements of the instructor. Be aware that instructors' policies on the granting of incomplete grades may vary and that the final decision on an incomplete grade is up to the instructor. **Instructors may deny requests for an incomplete grade.** If the missing assignment(s) have not been submitted to the instructor within 30 days from the end of the term in which the course was offered, or the agreed upon due date, the grade entered will reflect the student's grade in the course for the work completed and the missing assignments receiving no credit toward the final grade.

Attendance/Late Work Policy
Attendance Policy

In each term, students enrolled in on-ground sections are expected to be on campus and attending class beginning with the first day of classes. Students in online sections are expected to log in and participate in class beginning with the first day of classes.

Students who join a class after the first day of class during the university add period, or who are approved for late registration by the instructor and the Graduate School of Engineering, are responsible for all coursework missed prior to enrolling. In the interest of students' success, the college does not support the arrival of students to class after the university add deadline. **Enrolled students who do not attend class during the first week of a semester risk being dropped from the course.**

In cases where an enrolled student cannot arrive to campus by the first day of class due to circumstances beyond their control, it is the student's responsibility to contact the instructor for approval and notify the Graduate School of Engineering.

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 'F' for that course. Course instructors are not expected to make accommodations and students are expected to inform their instructors of any absences in advance of the class. 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 and accommodations in accordance with the University's academic and global entry expectations. Students may be asked to share communications about class absences with their Academic Advisor. 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-gradadvising@northeastern.edu) to learn more about the Medical Leave of Absence. 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.

Course Evaluations
Student feedback on their learning experience is valuable and helps improve future courses. We encourage all students to complete the course evaluation surveys when they become available.

Surveys are distributed at both the midterm mark and the end of the term via email and are completely anonymous and confidential. Any questions about the surveys can be directed to mgen-programs@coe.northeastern.edu

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

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.

The following is a broad overview, but not an all-encompassing definition, of what constitutes a violation of academic integrity:

*Cheating***:** The University defines cheating as using or attempting to use unauthorized materials, information, or study aids in any academic exercise. When completing any academic assignment, a student shall rely on their own mastery of the subject.

*Fabrication***:** The University defines fabrication as falsification, misrepresentation, or invention of any information, data, or citation in an academic exercise.

*Plagiarism***:** The University defines plagiarism as using as one's own the words, ideas, data, code, or other original academic material of another without providing proper citation or attribution. Plagiarism can apply to any assignment, either final or drafted copies, and it can occur either accidentally or deliberately. Claiming that one has "forgotten" to document ideas or material taken from another source does not exempt one from plagiarizing.

Unauthorized Collaboration: The University defines unauthorized collaboration as instances when students submit individual academic works that are substantially similar to one another. While several students may have the same source material, any analysis, interpretation, or reporting of data required by an assignment must be each individual's independent work unless the instructor has explicitly granted permission for group work.

Participation in Academically Dishonest Activities: The University defines participation in academically dishonest activities as any action taken by a student with the intention of gaining an unfair advantage over other students.

Facilitating Academic Dishonesty: The University defines facilitating academic dishonesty as intentionally or knowingly helping or contributing to the violation of any provision of this policy.

Please visit https://osccr.sites.northeastern.edu/academic-integrity-policy/ to access the full academic integrity policy.

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/Disability Access Services (DAS)
Northeastern University and the Disability Access Services (DAS) 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, students must provide appropriate documentation as provided by the DAS office.

If the course is conducted in an on-ground (in-person) format, students are expected to attend class physically as scheduled. Professors are **not required to provide virtual attendance links** unless a student has documented accommodation approved by the **Disability Access Services (DAS) office** and their **Academic Advisor**. If a student requires accommodation for remote participation, they must submit a formal request through the **Disability Office** and coordinate with their **Academic Advisor** prior to the course start date.

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

Office of Global Services
As an F-1, J-1, or Study Permit student, you must meet certain obligations in order to maintain lawful nonimmigrant status. Maintaining status is necessary in order to retain eligibility for the benefits of F-1 or J-1 status, such as employment authorization and program extension, and can be crucial to a successful application for a change or adjustment of nonimmigrant status in the future. Failure to maintain your nonimmigrant status can result in serious problems with immigration and *could lead to deportation from the U.S. or Canada.*

Students must maintain on-ground presence throughout the academic term. At Northeastern, there are four different defined instructional methods: Traditional, Hybrid, Live Cast, and Online. Traditional, Hybrid, and Live Cast courses meet the Visas' on-ground presence requirements. **Online courses do not meet the Visas' on-ground presence requirements.**

Students enrolled in Summer courses should adhere to OGS guidelines on maintaining status during the Summer term.

For more information please visit, https://international.northeastern.edu/ogs/current-students/understanding-visa-requirements/guidelines-on-maintaining-status/

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

Outreach, Engagement, Belonging

Northeastern University is committed to fostering a community of belonging, which is essential to the advancement of Northeastern University's mission of teaching and research. Our university is stronger as a result of the varied backgrounds, experiences, and perspectives that all members of our global community bring to the pursuit of knowledge. Embracing this pluralism is not the work of one office, department, or academic unit. It is a shared responsibility that spans disciplines and boundaries. By harnessing the power of our differences, we will continue to light the path to bold new ideas and life-changing discoveries.

It is my intention that students from all backgrounds and perspectives will be well served by this course, and that the diverse experiences that students bring to this class will be viewed as an asset. I welcome individuals of all ages, backgrounds, beliefs, ethnicities, genders, gender identities, gender expressions, national origins, religious affiliations, sexual orientations, socioeconomic background, family education level, ability – and other visible and nonvisible differences. All members of this class are expected to contribute to a respectful, welcoming and belonging environment for every other member of the class. Your suggestions are encouraged and appreciated.

Please visit Belonging at Northeastern – Northeastern Provost for complete information.

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.

The Office for University Equity and Compliance (OUEC) leads Northeastern University's efforts in maintaining compliance with all federal, state, and provincial civil rights laws and prohibits discrimination within any of its programs, activities, and services. Please visit https://ouec.northeastern.edu/ for more information and for the link to file a report.