Course Syllabus

Course Information

Course Title: ALY6040: Data Mining
Course Number (CRN) 70604
Term and Year: Summer B 2021
Start and End Dates:
9/20-10/30
Credit Hours: 3 Cr hrs
Course Format
Location: Hayden Hall 424
Meeting Days/Times: Thursdays 5:50pm

Instructor Information

Full Name: Justin Grosz
NEU Email Address: j.grosz@northeastern.edu

NuFLEX Requirements

Should this course fall under the Hybrid NuFlex policy, please refer to any and all supplemental materials pertaining to class attendance, participation, and other aspects impacting student and/or instructor engagement. For additional information, please refer to your advisor.

Technical Requirements

Courses are available on Northeastern University's Canvas at the following link:
http://canvas.northeastern.edu. Canvas Technical support and resources including 24/7 phone (1-833-450-3937), and chat can be found on the help icon in Canvas. Northeastern Technical support can be accessed at 617-373-4357 (xHELP) or help@northeastern.edu.

Each student is responsible for their access to the internet for purposes of this course and for research. Internet access is a required component of this course and will not be accepted as an excuse for missed work. If you know that you will be traveling, then make sure you plan accordingly.

Note regarding e-mail/voicemail: If you e-mail, please include your name and class title. Please allow up to 48 hours for an email reply. If you leave a voicemail, please remember to include your name, class title, and phone number.

Each student must be able to run R and RStudio on their own computers.

Course Prerequisites
ALY6000, ALY6010

Course Description
Introduces the theories and tools for intensive data analysis methods and data mining techniques such as rule-based learning, decision trees, clustering, and association-rule mining. Also covers interpretation of the mined patterns using visualization techniques. Offers students an opportunity to gain the knowledge and experience to apply modern data-mining techniques for effective large-scale data pattern recognition and insight discovery. Data analysis software is introduced. Student teams evaluate, analyze, and report data for the methods used and insights discovered during case studies.

Course Materials
- Exploratory Data Analysis with R (https://leanpub.com/exdata) - Free download
- rStudio software (https://www.rstudio.com/products/rstudio/download/#download) - Free download

- R software (https://cran.r-project.org/) - Free download

Python

- Jupyter Notebook Foundations: https://realpython.com/jupyter-notebook-introduction/
- Understanding Python Data Structure With Pandas: http://www.gregreda.com/2013/10/26/intro-to-pandas-data-structures/ (Links to an external site.)
- More Pandas But Within Jupyter Notebook: https://nbviewer.jupyter.org/github/fonnesbeck/Bios8366/blob/master/notebooks/Section2_1-Introduction-to-Pandas.ipynb (Links to an external site.)
- Bamboolib (GUI for Jupyter Notebook): Way to learn Pandas with GUI. https://bamboolib.8080labs.com/

Optional Materials

- Journal of Statistical Software (https://www.jstatsoft.org/index)

Program Learning Outcomes (PLOs)

PLO1: Demonstrate the foundational knowledge and skills critical to pursue data analytics as a profession in relation to statistics and math.

PLO2: Articulate and effectively defend the significance and implications of the work in data analytics in terms of challenges and trends in a local, national or global context.

PLO3: Demonstrate the knowledge of advanced tools in data analytics.

PLO5: Apply the principles, tools and methods of analytics to a comprehensive real-world problem or project related to data analyses for tactical and/or strategic decision making.

PLO6: Integrate the major theories, tools, and approaches in data analytics to identify data-driven insights for informed business process management.

SAIL Baseline Mapping

Enter 5 for Central, 4 for Significant, 3 for Moderate, 2 for Minimal, 1 for Potential, or 0 for None					*Enter 1 for Passive Engagement, 2 for Active Engagement, or 3 for Generative Engagement*
Social Consciousness & Commitment	**Global Mindset**	**Intellectual Agility**	**Personal & Professional Effectiveness**	**Well-Being**	**Level of Engagement**
2	2	5	3	1	3

Refer to SAIL for Web at https://sail.northeastern.edu/about/

Course Learning Outcomes

Based on satisfactory completion of this course, a student should be able to:

CLO1: Advocate for the use of data mining techniques to the leadership of an organization to achieve stakeholder alignment and business objectives

CLO2: Use the business objectives to discern "sound data" from noise with which to begin modeling to answer the relevant business questions

CLO3: Use data mining, EDA, and storytelling in to address a business problem in a way that conveys actionable insights to your peers and leadership

CLO4: Apply skills and knowledge of the analytics field to promote a data driven culture within an organization

CLO5: Choose, with appropriate reasoning, models and algorithms to complete a data modeling project

Expectations (Workload)

o   One (1) academic credit requires 50 minutes a week of classroom or faculty instruction and about two hours of out of class student work for a 15-week course; 100 minutes a week of classroom or direct faculty instruction and about 3.5 hours of out of class student work for a 7.5-week course.

o   For a three-credit course, students should expect 2.5 hours a week of classroom or faculty instruction and a minimum of 5 hours of out of class student work for a 15-week course; 5 hours of classroom or direct faculty instruction and a minimum of 10 hours of out of class student work for a 7.5-week course.

o   APA citations

Attendance Policy

Policy on late work

All assignments are to be turned in by deadline. If assignment is not received by deadline it is penalized 5 points per day. This will stand for all situations unless situation is addressed at least 24 hours before the deadline.

Course Methodology

Statements on how course instruction will be facilitated and expectations around participation such as presentations or discussion boards.

Participation/Discussion Board [ALL STUDENTS – ALL COURSE DELIVERY MODES]

- At least 1 **Primary responses** are due by 11:59 pm EST on the **Thursday** of each week
- At least 2 **secondary response** are due by 11:59 pm EST on the **Sunday** of each week

To facilitate interaction, students are expected to review the online postings on a regular basis even after they have posted their own minimum required postings. Please treat your classmates and the instructors with the utmost respect. Inappropriate posts will be removed immediately. The instructor reserves the right to penalize students for repeated violations of the participation policy (and/or Academic Integrity Policy) within a course. In the discussion board and in class, high quality contributions advance the class discussions and do not simply summarize the material that was assigned. Quality contributions take into account not only the instructor's questions but also your classmates' contributions. Please be mindful that the Discussion Board is a space for academic exchanges. As a result, students are accountable for using proper and exacting punctuation, spelling, and grammar. In addition, you may be required to reference all outside sources in correct citation format. It is crucial that all participants maintain a high regard for proper decorum in the Discussion Board.

Evaluation Standards

Rubrics and assignments points are outlined in each assignment in Canvas.

Grading

Graduate Programs Final Grading Scale

	87-89.9%	B+	77-79.9%	C+	

95-100%	A	84-86.9%	B	74-76.9%	C	69.9% or below	F

90-94.9%	A-	80-83.9%	B-	70-73.9%	C-	

Grade Breakdown:

Activity/Assignment Type	# Activities	Points Per Activity	This group's % of final grade
Discussions	6	100	15%
Module assignments	5	100	35%
Final Project and milestones • EDA • Project Draft • Final Project Report • Presentation	4	100	40%
Individual Reflection	1	100	10%
Total Course Points			100%

Use a separate Assignment Template to include details on the signature or other major assignments to clarify expectations such as purpose, task, resources, format, and rubrics.

Course Schedule

Week/Dates	Topic	Reading	Exercises/Assignments with due date
Week 1	Exploratory Data Analysis		Discussion Assignment 1: EDA assigned Final Project Proposal: Group assignments and data selection.
Week 2	Introduction to Data Mining Techniques		Discussion Assignment 1: EDA due Assignment 2: Technique Practice Final Project: EDA assigned

Week 3	Clusters, Association Mining, and Linear Discriminant Analysis		Discussion Assignment 3: Technique Practice Final Project: EDA due
Week 4	Support Vector Machines		Discussion Assignment 5: Technique Practice Final Project: Draft Report due
Week 5	Text Mining		Discussion Assignment 4: Technique Practice Final Project: Draft Report assigned
Week 6	Final Project & Wrap-up		Final Project Report submitted Final Project Presentation Individual Reflection

Notes to Instructor:

Exercises/Assignments with due date: If students need to submit an assignment by a certain time of day, please include time zone information.

End-of-Course Evaluation Surveys

Your feedback regarding your educational experience in this class is very important to the College of Professional Studies. 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 HuskyMail 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.

CPS expects independent and original work. Dishonesty and/or carelessness violates fundamental values of an intellectual and professional community and will be reported to the Office of Student Conduct and Conflict Resolution (OSCCR) as a violation of the Academic Integrity Policy. To safeguard the integrity of assignments and programs, your course may use systems such as TurnItIn, which checks written work, and Examity, an online exam proctoring system. Work that contains academic integrity violations (AIV) will be graded on a case by case basis. Work displaying AIV may earn a failing grade, a zero, or even result in the instructor assigning the student a failing grade for the class. Students may not withdraw from a class to avoid the grade penalty for serious academic integrity violations.

For additional information or for answers to frequently asked questions, please visit:

https://cps.northeastern.edu/academic-resources/academic-integrity

You may also email CPS-AcademicIntegrity@northeastern.edu.

Student Accommodations
The College of Professional Studies is committed to providing equitable access to learning opportunities to students with documented disabilities (e.g. mental health, attentional, learning, chronic health, sensory, or physical). To ensure access to this class, and program, please contact The Disability Resource Center (http://www.northeastern.edu/drc/) to engage in a confidential conversation about the process for requesting reasonable accommodations in the classroom and clinical or lab settings. Accommodations are not provided retroactively so students are encouraged to register with the Disability Resource Center (DRC) as soon as they begin their program. The College of Professional Studies encourages students to access all resources available through the DRC for consistent support.

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, visit http://library.northeastern.edu/.

Wyzant Tutoring Services
Northeastern University's College of Professional Studies is covering the cost of 4 hours of tutoring from Wyzant to help ensure your success in challenging courses. Tutors on Wyzant have helped thousands of students in 300+ subjects ranging from College Algebra to Academic Writing, Statistical Analysis to Microbiology. Click here to access your free tutoring through Northeastern University's Wyzant account.

Access your free tutoring using your Northeastern login credentials. Indicate your program, time zone, the course you'd like support in, and the specific topics or materials you'd like help with.

You'll receive an activation email from Wyzant for Higher Ed. (If you don't see it, check your spam or junk folders.) From the email, click on the activation button.
Potential tutors will begin reaching out to you. Their messages will be sent to your Northeastern email inbox and will be accessible via your Wyzant account dashboard.
 Respond back to the tutors you believe might be a good fit to schedule an online session when it's convenient for you — even late at night!

24/7 Canvas Technical Help
For immediate technical support for Canvas, call 1-833-450-3937. You can chat and report an issue to Canvas technical support right from within Canvas through the Help icon located in the Canvas global navigation.

myNortheastern, e-mail, and basic technical support
Visit the Information Technology Services (ITS) Support Portal
Email: help@northeastern.edu
ITS Customer Service Desk: 617-373-4357 (help)

Catalog
The College of Professional Studies Undergraduate Catalog is a reference/resource with information about curricula, resources, and academic and student policies.
For more information, visit http://www.cps.neu.edu/student-resources/.

Diversity and Inclusion
Northeastern University and the and the Office of Institutional Diversity and Inclusion (OIDI) is committed to equal opportunity, affirmative action, diversity and social justice while building a climate of inclusion on and beyond campus. In the classroom, member 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 sex and gender-based discrimination, including sexual harassment, sexual assault, sexual exploitation, relationship or domestic violence, and stalking. The Title IX Policy refers to sex and gender-based discrimination as "Prohibited Offenses." The Title IX Policy applies to the entire Northeastern community, including students, faculty and staff of all gender identities.

Allegations of Prohibited Offenses can be reported to the Title IX Coordinator within **The Office for University Equity and Compliance** at: titleix@northeastern.edu and/or through **the Northeastern University Police Department (NUPD)** by phone: for an Emergency 617.373.3333; for Non-Emergency 617.373.2121. Reporting to NUPD does **NOT** commit the victim/affected party to future legal action.

Faculty membe rs are considered "responsible employees"at Northeastern University, meaning they are required to report all reports of and information about alleged Prohibited Offenses to the Office for University Equity and Compliance.

If you or someone you know has experienced a Prohibited Offense, *confidential* support and guidance can be found through (https://www.northeastern.edu/ouec/resources/main-campus-resources/), **University Health and Counseling Services** (UHCS) staff (http://www.northeastern.edu/uhcs/) and the **Center for Spiritual Dialogue and Service** (CSDS) **clergy members** (http://www.northeastern.edu/spirituallife/). Employees within the VRC, UHCS, and CSDS are not required to report allegations of Prohibited Offenses to the Officefor University Equity and Compliance.

In case of an emergency, please call 911 OR NUPD'S Emergencyline: 617-373-3333.

 Please visit www.northeastern.edu/titleix for a complete list of reporting options and resources, both on-and off-campus.

The instructor reserves the right to amend this syllabus, both online and the document itself, during the term and will notify students of the change(s). The revised syllabus is the official record of class policies and schedule of due dates