CSCI 39543 – Intro to Data Mining

TuTh 7:00-8:15pm

Instructor: Justin Tojeira Semester: Spring 2021

Text: *Data Mining: The Textbook*. Charu C. Aggarwal. Springer, 2015. ISBN 978-3-319-14141-1

Instructor Email: jtojeira@hunter.cuny.edu

Course Description:
Intro to Data Mining will introduce students to the entire data mining process, including data collection, data preparation, and data mining algorithms. We will focus on data preparation. This involves understanding different types of data and the fundamental data mining tasks so that we may effectively select and prepare data as input to a data mining algorithm. Since there are countless algorithms designed for fundamental data mining tasks, many with multiple variations, we will only study the few most basic and most widely-used algorithms for each task. We will then apply the knowledge gained in a final project. I hope to cover all the material in the first 2/3 or so of the class and leave the last month of class to concentrate exclusively on the final project, though it remains to be seen if we will make it through the material in that time.

Online Classes:
Classes will be held every Tuesday and Thursday synchronously on Zoom. The Zoom link may be found on Blackboard in the Course Information folder.

Grading:
Because this is the first time Hunter College is offering this course, I must remain flexible in the material we will cover so that I assess our progress during the semester and adjust accordingly. Thus I cannot say with certainty if we will have 1 or 2 exams, or how extensive the final project will be.

1-2 midterm exams:	30 - 50%
Final Project:	20 - 50%
Homework:	20 - 30%

Final Project:
I would like the final project to be something that is portfolio-worthy, so I am aiming to finish the lecture portion of the class with a month remaining and have group projects involving collecting data from multiple sources, cleaning and agreggating it, selecting a sensible algorithm for your task, extracting features to feed into that algorithm, assessing your results, going back and adjusting your feature extraction to try to get better results as necessary, then vizualizing and presenting your results. However, whether we will be able to have a project with this scope depends on our progress in getting through our lecture topics.
During the time allocated to the final project, I will be assessing each group's progress and each student's contributions during each class session.

Homework:
Homework will be given on a regular basis, and we will go over it in class. Homework must be submitted on Gradescope by 6:55pm on its due date, which is right before class begins. Each homework will be graded either 1, representing that you have completed the homework or at least made a legitimate attempt, .5 representing that you left the homework largely unfinished or show a significant lack of effort, or 0 representing not submitting the homework or showing very minimal effort.
You will be allowed to make up 2 points worth of homework with late submissions. Your final homework grade will be normalized, e.g. if I decide that homework is worth 25 points of your final grade and give 10 homework assignments, each homework point will be worth 2.5 points on your final grade.

Blackboard (http://bb.hunter.cuny.edu/) will be used to post announcements, homework, exam dates, reviews for exams, and some material not found in the textbook. You should check Blackboard on a regular basis.

Additional Contact Information:
Computer Science Department: Room N-1008
Phone: 212-772-5213
Course Coordinator for CSCI 265: Eric Schweitzer, Room 1000E Hunter North
Phone: 212-772-4349
email: eric.schweitzer@hunter.cuny.edu

Hunter College Official Policies:
Hunter College regards acts of academic dishonesty (e.g., plagiarism, cheating on examinations, obtaining unfair advantage, and falsification of records and official documents) as serious offenses against the values of intellectual honesty. The college is committed to enforcing the CUNY policy on Academic Integrity and will pursue cases of academic dishonesty according to the Hunter College Academic Integrity Procedures.

In compliance with the American Disability Act of 1990 (ADA) and with Section 504 of the Rehabilitation Act of 1973, Hunter College is committed to ensuring educational parity and accommodations for all students with documented disabilities and/or medical conditions. It is recommended that all students with documented disabilities (Emotional, Medical, Physical and/or Learning) consult the Office of AccessABILITY located in Room E1124 to secure necessary academic accommodations. For further information and assistance please call (212-772-4857)/TTY (212-650-3230)