CSE 40647/60647 Data Science (Spring 2018) – Syllabus

Tuesday and Thursday, 2:00 pm to 3:15 pm, 117 DeBartolo

Instructor:

Dr. Meng Jiang, mjiang2@nd.edu
Office: 326C Cushing Hall Phone: (574) 631-7454

Teaching Assistant (TA):

Qi Li, qli8@nd.edu
Office: 212 Cushing Hall

Office Hours:

Instructor: **Friday**, 3:30 pm to 4:30 pm, 326C Cushing Hall
TA: **Monday**, 3:30 pm to 4:30 pm, 212 Cushing Hall
https://piazza.com/class/jcan36klo2c1e9
Please email me (mjiang2@nd.edu) to schedule any appointment outside of the office hours.

Course Webpage:

http://www.meng-jiang.com/teaching-cse647-s18.html

Text Book (not mandatory to have it):

- Jiawei Han, Micheline Kamber, and Jian Pei, "Data Mining: Concepts and Techniques," Morgan Kaufmann Publishers.
- We provide lecture notes from the 2nd edition of the text book on course website.

Prerequisites:

Programming with Python
Data structures and algorithms

As a prerequisite for:

CSE 40625/60625: Machine Learning

Course Goals:

At the end of the course, students will be able to:

- **Use** raw data processing techniques: data description, data visualization, data cleaning, data integration, data reduction, and dimension reduction
- **Use** Decision Trees, Naïve Bayes, and SVMs for classification
- **Describe** Ensembles and Neural Networks models for classification

- **Use** K-Partitioning methods for clustering
- **Describe** hierarchical clustering, kernel-based clustering and density-based clustering
- **Use** Apriori and FP-Growth for frequent pattern mining and association mining
- **Describe** diverse patterns, sequential patterns, graph patterns
- **Use** appropriate measures to evaluate results of different functionalities (classification, clustering, and frequent pattern mining)

Grading:

- Homework assignments: 20% = 5%*4
- Course project: 30%
- Mid-term exam: 20%
- Final exam: 30%

Any discrepancy in the grade should be brought to the attention of the instructor within **7** days of grade assignment.

Letter Grades:

- A: [93, 100]
- A-: [90, 93)
- B+: [87, 90)
- B: [84, 87)
- B-: [81, 84)
- C+: [78, 81)
- C: [75, 78)

Assignments:

- The assignments will require **individual effort**.
- The assignments will comprise of **paper-based and computer exercises** (i.e., **written/programming** assignments).

Course Project:

- The project will require **team effort**.
- The students will be required to write a **project term paper**.
- The project will be evaluated on **proposal, milestone presentation, oral presentation performance, and term paper quality**.
- Programming language will be **Python**.

Course Policies:

- Collaboration Policy

Unless instructed otherwise, students must turn in work that is their own. Students must write their own code, run their own data analyses, and write up their own results and answers to assignment questions.

- Assignments

There will be several assignments. Unless announced otherwise, assignments will be due at 11:59pm Eastern Time on the provided submission date.

- Quizzes and Exams

There will be no quizzes. There will be one mid-term exam and one final exam. Make-up exams will be allowed as per the du Lac Class Absence Policy. Whenever possible, students are expected to provide advance notice if they will be unable to take an exam. Make-up exams for travel to academic events will be provided at the discretion of the instructor.
Students are allowed to use only calculator but not any other electronic equipment (laptop, phone, etc.). Students are allowed to have one-page double-sided reference for the exams but not more than one letter-size piece of paper.

- Course Project

For the course project, students will be expected to follow the instructions, to propose a project topic and write a proposal paper, to give a milestone presentation and write a milestone paper, to write a term paper, and to present in class.

- Late Policy

Assignments submitted after the submission deadline but within the next day are counted as one day late. The next 24 hours will be counted as two days late, and so on. Each day late contributes a 33% penalty to the original assignment value.

- Office Hours

Students are encouraged to take advantage of the instructor and TA's office hours, and schedule additional time as needed.

Academic Dishonesty:

- The CSE and du lac honor code will be strictly followed.
- All assignments are individual unless instructed. You can discuss the assignment at a high level, but you should independently and individually write down the answers and/or the program. The sharing and copying of homework solutions or programs or functions or exams will be considered cheating.
- All the references and sources should be carefully provided and cited.
- Entering Notre Dame you were required to study the on-line edition of the Academic Code of Honor, to pass a quiz on it, and to sign a pledge to abide by it. The full Code and a Student Guide to the Academic code of Honor are available at: http://honorcode.nd.edu.
- Perhaps the most fundamental sentence is the beginning of section IV-B: "The pledge to uphold the Academic Code of Honor includes an understanding that a student's submitted work, graded or ungraded – examinations, draft copies, papers, homework assignments, extra credit work, etc. - must be his or her own."