

# **Project Instruction**

## **“Open Data Science Course Project”**

### **Data Science CSE 40647/60647**

Last updated: Aug 30, 2017 (being updated)

#### **Professor-in-charge:**

Dr. Meng Jiang, [mjiang2@nd.edu](mailto:mjiang2@nd.edu)

Office: 326C Cushing Hall

Phone: (574) 631-7454

#### **Teaching Assistant (TA):**

Qi Li, [qli8@nd.edu](mailto:qli8@nd.edu)

Office: 212 Cushing Hall

#### **Project goal:**

For the course project, students will be expected to collect a dataset (online or otherwise), formulate a question of interest, and perform aspects of data science to address that question by using whatever tools they find appropriate. The project will involve a proposal, milestones, and final term paper with a video presentation of the project.

#### **Project introduction:**

- The class project will require a proposal and milestones assessment during the semester with respect to the data science process (which will be provided in the first few weeks).
- The class project may involve some or all stages of the knowledge discovery process, depending on the chosen project.
- The students will be required to write a term paper and make a class presentation on their project.
- The term paper will go through a peer review process among the classmates.
- All project topics should be preapproved by the professor.
- The students may work in team of 2 – 4 (minimum 2 members are required) for the class project.

#### **Grading policy: (25% of the final score)**

Students are required to submit their data and code package + “readme” (.ZIP), project report/paper (.PDF), and video presentation (.MP4). There is no paper template requirement.

Students are encouraged to implement algorithms such as Apriori, FP-Growth, and K-Means Clustering by themselves instead of calling Python packages. Students are also encouraged to use Python packages (e.g., Numpy and Scipy) when they use advanced techniques (e.g., SVMs, Neural Networks, word2vec) to address challenging tasks.

Graders should have higher expectations on graduates than undergraduates – not only on the project results but also on writing (a workshop-quality paper of strong reasoning). Undergraduates will be applied with a uniform grading policy no matter what majors they have.

The project due is Nov 30, 2017. There will be NO extension. Significant updates are welcome before the final exam – students can send the updates to the instructor after the due by e-mail but they have to submit one version before the due.

### **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.”