

**CS 488/508: Introduction to Data Mining**

**Team Project**  
**Final Report and**  
**Presentation**

Tuan Le

College of Arts and Sciences

Computer Science



**BE BOLD.** Shape the Future.

# Final Report Requirements

- Deadline: **Oct 29, 11:55pm**
- These items should be discussed in the **final report**:
  - Data
    - What kind of data?
    - Dataset information such as number of instances, number of attributes, type of attributes
    - How do you process your data?
    - Which attributes you use and which one you don't use? Why?
  - Data mining task
    - What task? Classification, Clustering, Anomaly detection,...
  - Implementation and experiment results
    - Which algorithms you have tried?
    - How are the experiment results (e.g., accuracies, running time)?
  - Reference
- Submit **source code and slides. Please zip everything together and submit.**

# Final Presentation

- When: **Dec 3 and 5**
- Each team has **15 mins** (including Q&A)
- Talk about motivation, problem introduction, problem formulation (data mining task, data type), algorithms, experiment results, demo if you want,...
- Please submit your slides by **Oct 29, 11:55pm**

# Presentation Order

- This order is randomly generated:
  1. MedINFOExtract (TL: Hiep)
  2. FUMP (TL: Aaron Saucedo)
  3. Team 1 (TL: Ziwei Ma)
  4. Data Prospectors (TL: Jorge Garcia)
  5. Team 4 (TL: Ivan Nieto)
  6. SSN (TL: Sanuj Kumar)
- Starting from **9am on Dec 3**