

2017-2018

Capstone Projects Introduction

DSE 260A (Winter) and 260B (Spring)
MAS on Data Science and Engineering

Dr. Ilkay Altintas
altintas@sdsc.edu

What is a Capstone Project?

- **Objective:** To complete an end to end analysis of a large (>10GB) dataset with big data characteristics.
 - Includes
 - data collection,
 - ETL,
 - exploratory analysis,
 - model building, and
 - visualization / reporting.
- **Products:**
 - Final report (preferred if publishable as a conference paper)
 - Output data products
 - Developed analytical tools/methods/workflows (if applicable)

Milestones for the Capstone Project

- **First Year:**
 - Spring Quarter: Capstone project class is introduced.
 - Summer: Advisors provide short summaries of projects so that students can identify who they want to work with. Students start to form teams, define project and find advisor. (Please wait for email later this summer from capstone faculty, i.e., Altintas.)
- **Second Year:**
 - Fall Quarter: Students finalize project teams and start collecting/exploring data.
 - Winter Quarters: Teams work on their projects and present monthly progress reports. Alternate meetings: once a month for 2 hours with advisor, twice a month for half an hour with capstone faculty (i.e. Altintas).
 - Spring quarter: Teams finalize their projects, including documentation and final report. Teams make open presentations to their peers, advisor and capstone faculty, and receive final grade.

Question: Is team work allowed?

- Team work is required.
- A project team can consist of 2-5 students
- A full project plan on distribution of tasks and roles should be submitted to and approved by the advisor and Dr. Altintas

Question: How much work per week is expected?

- A project should require 5-10 hours per week by all project participants.
- The activities per week should be discussed in a regular meeting with advisors
 - Meetings at least twice month is advised

Question: Can I use company data or work on a company related project?

- Admissible under the following conditions:
 1. The input data and the results files will be made available to the project advisor.
 2. The final report on the project, on which the project grade will be based, is a public document.

2017-2018 Projects

Will be announced at:

<http://mas-dse.github.io/capstone/>

Next Step 1: Meet the advisors!

- Documents will be published on the website by potential advisors
 - Please wait for email on the final projects
<http://mas-dse.github.io/capstone/>
- Arrange meetings with the advisors to discuss their interests
- Email me (ialtintas@ucsd.edu) to ask if you have any questions
 - *** Start email subject with “MAS/DSE Capstone 17-18:” ***

Next Step 2: Think of a project

- Start early!!!
- Things to consider:
 - Data sources
 - Analysis problems
 - What MAS/DSE knowledge you will be applying
 - Feasibility in a year
 - Output organization

Next Step 3: Plan your Project

- Create a design document and milestones based on the first three steps
- Agree on team responsibilities
- Create a regular meeting calendar
- Identify potential road blocks

2016-2017 Projects

- Project 1: Traffic Data Analysis
 - Project 2: Wildfire Satellite Data Analysis
 - Project 3: Amazon Data Analysis
 - Project 4: Fitness Data Analysis
 - Project 5: Smart Building Data Analysis
-
- You are invited to the presentations on June 9th

What is a general process for a capstone project?

Questions?