

DSCI 560: Data Informatics Professional Practicum

32422D — Fall 2020

Instructors Information

Name: Daniel Garijo
Email: garijo@usc.edu
Office Hours: W - 4:30-5:30pm - Zoom: 950 2322 2216

Name: Deborah Khider
Email: khider@usc.edu
Office Hours: Tu - 4-5pm - Zoom: 980 1985 7991

Graders Information

Name: Avantika Shenoy
Email: avantika@usc.edu

Name: Pratishtha Singh
Email: pratisht@usc.edu

Class Information

Dates: 8/28/2020 – 12/9/2020
Time: 13:00 – 16:50
Classroom: Zoom: 990 0596 2565

This course will be held online synchronously (you MUST be logged in with your USC NetID to access the Zoom room) and asynchronously (recording will be made available through Blackboard). According to USC Viterbi School Engineering's policy, *class recordings are made available only for access by students in this class for educational purposes only and shall not be disclosed to any other party for any purpose.*

All class communication, materials and announcements will be made via Blackboard.

Course Description

Student teams working on external customer data analytic challenges; project/presentation based; real client data, and implementable solutions for delivery to actual stakeholders; capstone to degree. *Recommended preparation: Knowledge of data management, machine learning, data mining, and data visualization.*

Expanded Course Description

One of the fundamental principles of data science is the ability to live in and understand data. It is necessary to gain a level of immersion in the information environment to truly apply the diverse skill sets necessary to both become an effective analyst, and provide solutions to hard problems. The goal of this course is to combine previously learned capabilities in data management, machine learning, data visualization, and knowledge representation and apply them against actual

data sets, in real data environment, and toward solving difficult challenges. The Data Informatics Professional Practicum is a capstone experience designed to allow students exposure to the world of data analytics from the perspective of a stakeholder. Students will work with stakeholders on a project that answers a data analytic problem. Students will team up and propose projects around these requirements, and be responsible for delivering an informatics-based solution to the stakeholder. A possible example of a capstone project would be: a non-profit organization volunteers to be the external partner. They have a dataset about crimes in Los Angeles. The NGO might pose the question “How can the data we collect and process data to understand the overall sentiment toward climate change in the United States?” It would be up to the student teams to proposal a project plan, design a study, implement the analysis, and report conclusions and recommendations. The Professional Practicum (DSCI 560) provides an opportunity to face real-world challenges and to develop your personal innovative project in cooperation with the members of your team. The hands-on experience with an external partner will provide a peek at what to expect as a data informatics experts and data scientist, in addition to providing project management techniques.

Learning Objectives

Expected learning outcomes for students are:

- To understand how to apply the various engineering principles studied in Data Informatics curriculum toward solving an organizational challenge
- To handle difficulties associated with defining and organizing a realistic problem statement
- To manage impediments in obtaining information and approval
- To present and sell ideas to higher-level management
- To understand the importance of the need for a continuous exchange between engineers, management and employees in solving an existing problem, given a set of constraints
- To meet aggressive deadlines in a multidisciplinary team effort
- To improve project-based presentation skills
- To understand the requirements and objectives of customers, how these vary, and how one must tailor a solution to the expectations of a customer
- To understand how informatics are viewed in diverse domains; and how solutions and approaches will differ in various fields
- To understand how to work with individuals from diverse domain (engineering, science, etc.) to accomplish a common goal

Homework

There will be five homework assignments throughout the course. The homework assignments will test students understanding of the concepts learned in class but will be independent from the main class project. The assignments must be submitted individually and students will receive individual scores. Students may work in groups to complete the tasks. The homework assignments are expected to take 6-8 hours. Each assignment is graded on a scale of 0-100 and the grading criteria will be specified in each assignment. Homework assignments are due at 12:00pm Pacific on the due date and should be submitted via Blackboard. Homework will be accepted up to one week late as long as the student requested a late submission ahead of the deadline for a **documented** emergency (family emergency, illness), and in that case the assignment will be graded at 20% less than the possible points for the assignment. After one week, the assignment will not be graded.

Project

The main purpose of this course is to complete a data science project on external customer data. The final report should be no less than 50 pages detailing the proposal, planning, methods utilized in the analysis, the output from the analysis, and conclusions and recommendations based on the study. In addition, students are expected to share their data, software, and methods according to best practices discussed in class. This assignment must be completed within an assigned group.

The final project report will be due Dec 2nd, 2020 at 5:00pm Pacific. No late report will be accepted except in the case of a documented emergency for the entire group.

Presentation

Students will also be responsible for an in-depth 30 minute presentation on their project to their peers. The presentation will need to be recorded before the dress rehearsal and final presentation and uploaded to Blackboard at 5:00pm Pacific the Thursday before the presentation.

Grading

The course grade is determined by the following components:

Homeworks	500 points
Presentation	200 points
Project report	500 points

Grade Scale

Final grades will be assigned according to the following scale:

A	1128 – 1200	C+	924 – 959
A-	1080 – 1127	C	888 – 923
B+	1044 – 1079	C-	840 – 887
B	1008 – 1043	D	720 – 839
B-	960 – 1007	F	0 – 719

Tentative Schedule

The class will be divided into two periods: the first two hours will be devoted to a lecture. The last two hours will be hands-on exercises (except in the first two weeks of the semester). These hands-on exercises will be completed asynchronously with QA through Blackboard forums. Each hands-on exercise will have its own forum with a single question per thread. Instructors will answer questions through the forum on Fridays 4-5pm and Wednesdays 9-10am.

Week	Lecture Topic	In-class practicum	Homework
1	Syllabus Overview Learning to work in a data science team	Learning different situations when working as a team in data science projects	
2	Project scope	Introduction to the course data science projects	Submit resume Homework 1: Scoping your project
3	Data and metadata	Upload datasets to appropriate platform and describe metadata	Homework 1 due Homework 2: Data, software and workflows
4	Software and metadata	Upload software to appropriate platform and describe metadata	
5	Project Management (I)	30-min session with instructor about your group's project scope	Homework 2 due
6	Project Management (II)	30-min session with instructor about your group's project 1st iteration	Homework 3: Project delivery plan
7	Containers and reproducibility	Practicum with Docker	
8	Visualization	Create a simple visualization in Bokeh/ipywidgets	Homework 3 due Homework 4: Containerization
9	Data science communication	Giving talks to different audiences: Talk 1: Meeting with technical team (peers) Talk 2: Talking to managers Talk 3: Talking to stakeholders	Homework 5: Visualization
10	Guest Lecturer: Alex Remedios	Schedule 30min meeting with instructors to go over project	Homework 4 due
11		Schedule 30min meeting with instructors to go over project	Homework 5 due
12		Presentation dress rehearsals	
13		Project Presentations	

As time zone permits, meetings with instructors will be scheduled during the normal call period after the lecture (Fridays, 3-5pm).

Honor Code

In response to recommendations made by the Academic Integrity Task Force to the Dean, the USC Viterbi School of Engineering now has an Honor Code. The Code was developed by Viterbi students, and its text is as follows:

Engineering enables and empowers our ambitions and is integral to our identities. In the Viterbi community, accountability is reflected in all our endeavors.

Engineering+ Integrity. Engineering+ Responsibility. Engineering+ Community. Think good. Do better. Be great.

These are the pillars we stand upon as we address the challenges of society and enrich lives.

During your time here at Viterbi, please know that academic and personal resources are available to help you:

- The student-driven and student-written Honor Code is here: <http://viterbi.usc.edu/academics/integrity/>.
- An introductory video is posted at <https://myviterbi.usc.edu/> under the link "Academic Integrity Introduction" and serves as a reminder of the school's emphasis in maintaining a high level of academic integrity.
- Master's and PhD students can contact the GAPP office in OHE 106 (<https://gapp.usc.edu/>) for other helpful resources.
- The Viterbi Academic and Resource Center (VARC) (<http://viterbi.usc.edu/students/undergrad/varc>) has a variety of services available.

Academic Integrity

The Viterbi School takes academic integrity violations seriously. Most of the violations that have been reported in the past fall into four categories: unauthorized collaboration, plagiarism, code sharing, and cheating on an exam. Specifically:

- Unauthorized collaboration - Unauthorized collaboration on a project, homework or other assignment. (section 11.14.B) All homework assignments must be individually developed. Students that collaborate on assignments will be referred to the Academic Integrity Coordinator.
- Plagiarism - presenting someone else's ideas as your own, either verbatim or recast in your own words - is a serious academic offense with serious consequences.
- Code sharing - Obtaining for oneself or providing for another person a solution to homework, a project or other assignment, without the knowledge and expressed consent of the instructor. (section 11.14.A)
- Cheating in an exam - this may involve a number of violations, such as looking at class notes during the exam, looking at other student's exam, "texting" with other students during the exam. See the section titled Two Exams for a list of specific violations.

Please note that that these are only the basic violations that we have encountered in the past, and there are many more. Please familiarize yourself with the discussion of plagiarism in SCampus in Section B.11.00, Behavior Violating University Standards and Appropriate Sanctions available at <https://scampus.usc.edu/b/11-00-behavior-violating-university-standards-and-appropriate-sanctions/>. All academic integrity violations will be referred to the Academic Integrity Coordinator of the Viterbi School of Engineering. The process for adjudicating these cases is available in SCampus, Part B, Section 13.

Other Misconduct

Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, <http://policy.usc.edu/scientific-misconduct/>.

Discrimination, sexual assault, and harassment are not tolerated by the university. You are encouraged to report any incidents to the Office of Equity and Diversity <http://equity.usc.edu/> or to the Department of Public Safety <http://capsnet.usc.edu/department/department-public-safety/online-forms/contact-us>. This is important for the safety whole USC community. Another member

of the university community - such as a friend, classmate, advisor, or faculty member - can help initiate the report, or can initiate the report on behalf of another person. The Center for Women and Men <http://www.usc.edu/student-affairs/cwm/> provides 24/7 confidential support, and the sexual assault resource center webpage <http://sarc.usc.edu> describes reporting options and other resources.

Support Systems

A number of USC's schools provide support for students who need help with scholarly writing. Check with your advisor or program staff to find out more. Students whose primary language is not English should check with the American Language Institute <http://dornsife.usc.edu/ali> which sponsors courses and workshops specifically for international graduate students. The Office of Disability Services and Programs http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html provides certification for students with disabilities and helps arrange the relevant accommodations. If an officially declared emergency makes travel to campus infeasible, USC Emergency Information <http://emergency.usc.edu/> will provide safety and other updates, including ways in which instruction will be continued by means of blackboard, teleconferencing, and other technology.

Diversity

The diversity of the participants in this course is a valuable source of ideas, problem solving strategies, and engineering creativity. The instructors encourage and support the efforts of all of our students to contribute freely and enthusiastically. As members of an academic community, it is our shared responsibility to cultivate a climate where all students and individuals are valued and where both they and their ideas are treated with respect, regardless of their differences, visible or invisible.

Students with disabilities

Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to TA) as early in the semester as possible. DSP is located in STU 301 and is open 8:30 a.m. - 5:00 p.m., Monday through Friday. Website and contact information for DSP: http://sait.usc.edu/academicsupport/centerprograms/dsp/home_index.html, (213) 740-0776 (Phone), (213) 740-6948 (TDD only), (213) 740-8216 (FAX), ability@usc.edu.

Emergency Preparedness/Course Continuity in a Crisis

In case of a declared emergency if travel to campus is not feasible, USC executive leadership will announce an electronic way for instructors to teach students in their residence halls or homes using a combination of Blackboard, teleconferencing, and other technologies.