**SSCI 575, Spatial Data Science**

*Syllabus*

**Units:** 4

**Term Day Time:** Tue, Thu 10:00 - 11:50 am

**Location:** AHF 145A and DEN@Dornsife

**Instructor:** Yi Qi, PhD
**Office:** AHF B55J
**Office Hours:** Thu, 1:00 – 3:00 pm, or by appointment
**Contact Info:** yi.qi@usc.edu, 213-821-1589

**Library Help:** Andy Rutkowski
**Office:** LIPA B40-A
**Office Hours:** Thu 10am - 12 pm or by appointment
**Contact Info:** arutkows@usc.edu

**IT Help:** Myron Medulla
**Office:** AHF B56B
**Office Hours:** By appointment via email
**Contact Info:** spatial_support@usc.edu, 213-821-4415

## Course Scope and Purpose

This course is a required course for the Master's Degree in Spatial Data Science, and an elective course for the Master's Degree in Geographic Information Science & Technology (GIST), Master's Degree in Spatial Economics and Data Analysis, Graduate Certificate in Geospatial Leadership, Graduate Certificate in Remote Sensing for Earth Observation (RSEO), and Transportation Systems Management GIS Specialization.

Could we have prepared for an earthquake evacuation using data such as geotagged tweets? How are our movement patterns related to spreading of a pandemic? Questions such as these are becoming more commonplace in spatial analysis as data sources increase in variety, and the questions we ask are more interdisciplinary. This course aims to provide students with theoretical and practical knowledge pertaining to spatial machine learning methods frequently used in spatial analysis. Students will gain experience in statistical, algorithmic and practical applications of modern spatial machine learning methods that complement and enhance traditional (spatial) statistical approaches to spatial analysis. Students will be introduced to statistical computing in the context of machine learning.

The course is structured as a series of theoretical and hands-on lectures, where students will learn theory and algorithms behind commonly used machine learning methods and apply them to spatial analysis problems. The course will use Geographic Information Systems (GIS) and Python as the main computational framework for representing spatial data and all class examples will be studied within the context of GIS. This course features a series of hands-on projects where students will use cutting-edge methods to solve a spatial data science problem that span multiple disciplines, such as public health, the house market, and natural disaster response, to name a few. Students will also perform research and reading assignments, where they will be studying state of the research in different aspects of spatial machine learning and making connections with research and foundational ideas covered in class.

This a graduate level course, so you should expect this class to be both academically robust and intellectually challenging. As graduate students you are expected to engage with the information you are learning and to explore the heady cauldron of ideas, opinion, and analysis that describe our collective effort to thoroughly interrogate the subject at hand.

### *Learning Outcomes*

On completion of this course, students should be able to:
- Identify different types of machine learning methods in terms of their formulation
- Apply appropriate type of machine learning methods to spatial problems
- Improve machine learning model performance through the use of diagnostics
- Explain differences between spatially implicit and explicit machine learning methods

- Perform medium to large scale analysis on imperfect field data for creating data driven models of social and natural phenomena
- Work with unconventional spatial data sources and integrate machine learning into their analysis pipeline
- Implement the spatial machine learning workflow for complex problems and systematically solve a multidisciplinary problem
- Describe algorithmic differences between different machine learning methods
- Perform exploratory analysis to delineate best methods for the data and problem at hand

Students may vary in their competency levels on these abilities. You can expect to acquire these abilities only if you honor all course policies, attend classes regularly, complete all assigned work in good faith and on time, and meet all other course expectations of you as a student.

**Prerequisite(s)**: None
**Co-Requisite(s)**: None
**Concurrent Enrollment:** None
**Recommended Preparation:** SSCI 581: *Concepts for Spatial Thinking*

## Class Conduct

**Harassment, sexual misconduct, interpersonal violence, and stalking** are not tolerated by the university. All faculty and most staff are considered Responsible Employees by the university and must forward all information they receive about these types of situations to the Title IX Coordinator. The Title IX Coordinator is responsible for assisting students with supportive accommodations, including academic accommodations, as well as investigating these incidents if the reporting student wants an investigation.  The Title IX office is also responsible for coordinating supportive measures for transgender and nonbinary students such as faculty notifications, and more.  If you need supportive accommodations you may contact the Title IX Coordinator directly ([titleix@usc.edu](mailto:titleix@usc.edu) or 213-821-8298) without sharing any personal information with me. If you would like to speak with a confidential counselor, Relationship and Sexual Violence Prevention Services (RSVP) provides 24/7 confidential support for students (213-740-9355 (WELL); press 0 after hours)

**COVID-19 policy** -- Students are expected to comply with all aspects of USC's COVID-19 policy including, but not limited to, vaccination, indoor mask mandate, and daily TrojanCheck. Failure to do so may result in removal from the class and referral to Student Judicial Affairs and Community Standards. Students are recommended to keep safe physical distancing, whenever possible, to prevent any possible transmission. Please contact your instructor if you have any safety concerns.

**Diversity and Inclusion** – It is my intent that students from all diverse backgrounds and perspectives be well served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that students bring to this class be viewed as a resource,

strength and benefit. It is my intent to present materials and activities that are respectful to everyone, and you are also expected to respect of others regardless of their race, ethnicity, gender identity and expressions, cultural beliefs, religion, sexual orientation, national origin, age, abilities, ideas and perspectives, or socioeconomic status. Your suggestions are encouraged and appreciated. Feel free to let me know ways to improve the effectiveness of the course for you personally or for other students.

## Course Structure

The main theoretical concepts are introduced in class via lectures, discussions and after class via reading assignments. Practical exercises utilize Python and the ArcGIS platform for problem-solving. Some assignments may require coding, and working knowledge of Python will be assumed.

*Workload* – This is a four credit, one semester course. Students should expect to spend 10-15 hours per week completing the work in this course.

## Technological and Communication Requirements

All software this course will use is provided online via the GIST Server; hence, you do not need to install it on your own computer. Instead, every student must have the following technology requirements:

- A computer with a fast Internet connection.
- A functional webcam and a microphone for use whenever a presentation or meeting is scheduled.
- An up-to-date web browser to access the Server

If a student does not have access to any of these, please speak with the instructor at the start of the semester. Also, see the USC ITS Student Toolkit here: https://keepteaching.usc.edu/students/student-toolkit/

*Desire2Learn (D2L)* – This course will utilize the Desire2Learn (D2L) learning management system which allows students to access course content, upload assignments, participate in discussion forms, among other learning experiences.  The D2L platform provides flexibility in the learning experience where students can participate in the course residentially or remotely, synchronously (meeting together at the same time) or asynchronously (accessing videos and course content outside of class).

*SSI Server and Tech Support* – This course utilizes the SSI Server which is a virtual desktop giving access to many different professional software. If you are unable to connect to the server or experience any type of technical issues, send an email using your USC account to SSI Tech Support at spatial_support@usc.edu, making sure to copy (cc) me on the email.

*Communications* – All assignments given and all materials to be handed in will be submitted via D2L. The instructor will also create and monitor discussion forums through which students can discuss issues and assignments as needed.  Students should read all email sent from D2L or from course instructor(s) as soon as possible.  Also, students who do not regularly use

their USC email accounts should double-check to be sure that mail sent from both the D2L accounts and the instructor's account (noted above) to your USC account is forwarded to an address used regularly and does not go into junk mail.  The instructor will endeavor to respond to all email within 24 hours of receipt, aiming for no more than 72 hours delay.  In the rare case that an instructor is off-line for an extended periodof time, an announcement will be posted to the class D2L site.  Due to the synchronous and asynchronous nature of this course, it is each student's responsibility to stay informed and connected with others in our course. In addition to email, students are expectedto login to D2L regularly to check for announcements.

*Discussion forums* – On the D2L site, I will post a series of discussion threads relevant to various sections of the course. Discussions provide a key means for student-to-student discussion and collaboration that can replicate the face-to-face contact you may have experienced in traditional classrooms. Here students can provide support to each other while working on your assignments, sharing hints and helpful tips, as you would in a classroom laboratory. Please post your questions about assignments there, as you would ask them publicly in the classroom. I monitor the discussion threads and offer comments when necessary, but more importantly, consider the discussion board a key way to connect with your classmates and share your discoveries.

## Required Readings and Supplementary Materials

The optional textbooks for this course are:

- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. Berlin, Germany: Springer Science & Business Media. Official Link to PDF

- Haining, R.P. (2003). Spatial data analysis: Theory and practice. Cambridge, UK: Cambridge University Press. (an online copy is available via USC library)

Assigned supplementary readings are below (asterisk (*) indicates optional reading):

- Janowicz, K., Gao, S., McKenzie, G., Hu, Y., & Bhaduri, B. (2019). GeoAI: Spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. International Journal of Geographical Information Science, 34(4), 625-636.

- Duque, J.C., Ramos, R., & Suriñach, J. (2007). Supervised regionalization methods: A survey. International Regional Science Review, 30(3), 195-220.

- Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, A. S. (2019). mgwr: A Python implementation of multiscale geographically weighted regression for investigating

process spatial heterogeneity and scale. ISPRS International Journal of Geo-Information, 8(6), 269.

- Battaglia, P.W., Hamrick, J. B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., ... Gulcehre, C. (2018). Relational inductive biases, deep learning, and graph networks. arXiv preprint, arXiv:1806.01261.

- Tong, D., & Murray, A.T. (2012). Spatial optimization in geography. Annals of the Association of American Geographers, 102(6), 1290-1309.

## Description and Assessment of Assignments

### *Weekly Assignments*

There are several different kinds of assignments with at least one due weekly. These are described in the Weekly Folders on D2L. Due dates are shown in the summary that follows.

*Resume Assignments – 1 worth a total of 5 points*. We require all current students to post and maintain a public resume, short biography and recent photo on our shared SSI Student D2L site. Please prepare your resume in the SSI template which will be provided to you. Unless you opt out, your resume will be included in the Spatial Sciences Institute Graduate Programs Resume Book. This resume book is compiled annually and, along with our web presence, is used to promote our programs, and more importantly, your skills, experience and professional aspirations.

*Projects – 3 worth a total of 30 points*. 3 projects and will be the major tool used to evaluate your learning in this course. The projects will be linked to course Modules. In support of these projects, students will complete end-to-end data science workflows. Students will gain insight to executing a data science workflow with clear deliverables.

*Reading Assignments – 5 worth a total of 25 points*. Each student is required to complete all reading assignments for this class. The reading assignments will focus on the theory portion of the course as presented in the weekly readings. The objective of the reading is to help you evaluate and integrate the information you have acquired from the course readings. Some of these will involve discussions and collaborative work and some will be individual efforts.

*Mid-Term Exam – 1 worth of a total of 10 points*. The mid-term will cover material learned in the first half of the semester. It may be mixed format and may consist of multiple choice, short answer, and hands-on calculations.

*Final Exam – 1 worth a total of 10 points*. The final exam will cover material learned over the duration of the semester. It may be mixed format and may consist of multiple choice, short answer, and hands-on calculations.

### *Final Project*

The final project will provide you the opportunity to integrate learning from various aspects of the course through the practice of a more in-depth assignment. The final project will require independent thoughts and synthesis. You will choose one of the two options to craft a written document: 1) report a case study where you design and complete a full chain of analysis using remotely sensed data; 2) conduct a literature review on one topic of your interests. More detailed instruction on the final project will be discussed in the class.

*Final Report - 1 worth 15 points* – The final report is a professional writing of your finished research project or literature review.

*Final Presentation - 1 worth 5 points* – You will present your final project in a 10-min oral presentation via a PowerPoint in Zoom in the week preceding the final exam.

## Grading Breakdown

Careful planning and a serious, consistent commitment will be required for you to successfully navigate the various deliverables in this course. The table below summarizes the SSCI 575 course assignments and their point distribution. There are 5 points designated to gauge your levels of engagement in class discussion with the professor and classmates, including but not limited to in-class discussion, D2L discussion forum, meetings with professors, and so forth.

| Assessment | Number | Points Each | Total Points |
|---|---|---|---|
| Resume Assignment | 1 | 2 | 2 |
| Projects | 3 | 10 | 30 |
| Final Report | 1 | 15 | 15 |
| Final Presentation | 1 | 5 | 5 |
| Reading Assignments | 5 | 5 | 25 |
| Mid-Term Exam | 1 | 10 | 10 |
| Final Exam | 1 | 10 | 10 |
| Class Discussion | 1 | 3 | 3 |
| Total | 14 | -- | 100 |

## Assignment Submission Policy

Unless otherwise noted, all assignments must be submitted via D2L *by 11:59 pm Pacific Time (PT) on due dates specified in* the Course Schedule below. Your attention to on-time assignment submission is essential if I am to meet my goal to return comments on your submitted assignments before the next one is due. Sometimes this is impossible, so I will post a notice on anticipated delays if needed.

Strict penalties apply for late assignments as follows:

- All assignments will be penalized 2 points up to FOUR days late. No points will be given for submissions more than FOUR days late. Note that all assignments worth 2 points will receive 0 points if submitted late.

- Additionally, no written work will be accepted for grading after 5 pm PT on the last day of classes.

Any exceptions to these turn-in assignments are only made by me in coordination with individual students. An example of an exception would be a student's illness or injury that reasonably prohibits course involvement/participation.

## Schedule

| | Topic | Readings and Assignments | Deliverables/Due Dates |
|---|---|---|---|
| **Module 1: Data Representation and Wrangling** | | | |
| **Week 1** 8/21 | Is Spatial Special? Introduction to Spatial Data Science and GeoAI | Resume Assignment | No deliverables |
| **Week 2** 8/28 | Introduction to the Spatial Data Science Workflow | Reading Assignment 1 | Resume Assignment |
| **Week 3** 9/5 *Monday, 9/4 is university holiday (Labor Day) | Spatial Data Representation Introduction to Spatial Statistics | Project 1 | |
| **Week 4** 9/11 | Spatial Data Wrangling | Reading Assignment 2 | Reading Assignment 1 |
| **Module 2: Exploratory Spatial Data Analysis** | | | |
| **Week 5** 9/18 | Exploratory Spatial Data Analysis | | |
| **Week 6** 9/25 | Pattern Detection | | Project 1 |
| **Module 3: Unsupervised Learning** | | | |
| **Week 7** 10/2 | Spatially Explicit Unsupervised Learning | Project 2 | Reading Assignment 2 |

| | Module 4: Supervised Learning | | |
|---|---|---|---|
| **Week 8** 10/9 *10/12-10/13 is university holiday (Fall Recess) | Spatially Explicit Linear Models | Reading Assignment 3 | |
| **Week 9** 10/16 | Spatially Explicit Machine Learning | | Mid-Term Exam |
| **Week 10** 10/23 | Spatially Explicit Neural Networks I | Project 3 | Project 2 |
| **Week 11** 10/30 | Spatially Explicit Neural Networks II | Reading Assignment 4 | Reading Assignment 3 |
| | **Module 5: Model Assessment** | | |
| **Week 12** 11/6 *Friday, 11/10 is university holiday (Veterans Day) | Model Assessment and Selection | Reading Assignment 5 | |
| | **Module 6: Spatial Optimization and Sentiment Analysis** | | |
| **Week 13** 11/13 | Spatial Optimization Sentiment Analysis in Spatial Analysis | | Project 3 |
| **Week 14** 11/20 *11/22-11/24 is a university holiday (Thanksgiving) | No class | | Reading Assignment 4 |
| **Week 15** 11/27 *Friday, 12/1 is the last day of | Course Review (11/29) Oral Presentation (12/1) | | Reading Assignment 5 Final Project |

| class, 12/2-12/5 study days | | | |
|---|---|---|---|
| **Final Exams** 12/6-12/13 | Final Exam | | |

## Statement on Academic Conduct and Support Systems

### *Academic Integrity*

The University of Southern California is a learning community committed to developing successful scholars and researchers dedicated to the pursuit of knowledge and the dissemination of ideas. Academic misconduct, which includes any act of dishonesty in the production or submission of academic work, comprises the integrity of the person who commits the act and can impugn the perceived integrity of the entire university community. It stands in opposition to the university's mission to research, educate, and contribute productively to our community and the world.

All students are expected to submit assignments that represent their own original work, and that have been prepared specifically for the course or section for which they have been submitted. You may not submit work written by others or "recycle" work prepared for other courses without obtaining written permission from the instructor(s).

Other violations of academic integrity include, but are not limited to, cheating, plagiarism, fabrication (e.g., falsifying data), collusion, knowingly assisting others in acts of academic dishonesty, and any act that gains or is intended to gain an unfair academic advantage.

The impact of academic dishonesty is far-reaching and is considered a serious offense against the university. All incidences of academic misconduct will be reported to the Office of Academic Integrity and could result in outcomes such as failure on the assignment, failure in the course, suspension, or even expulsion from the university.

For more information about academic integrity see the student handbook or the Office of Academic Integrity's website, and university policies on Research and Scholarship Misconduct.

Please ask your instructor if you are unsure what constitutes unauthorized assistance on an exam or assignment, or what information requires citation and/or attribution.

### Students and Disability Accommodations:

USC welcomes students with disabilities into all of the University's educational programs. The Office of Student Accessibility Services (OSAS) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to be reasonable and appropriate, a

Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

**Support Systems:**

*Counseling and Mental Health - (213) 740-9355 – 24/7 on call*

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

*988 Suicide and Crisis Lifeline - 988 for both calls and text messages – 24/7 on call*

The 988 Suicide and Crisis Lifeline (formerly known as the National Suicide Prevention Lifeline) provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week, across the United States. The Lifeline is comprised of a national network of over 200 local crisis centers, combining custom local care and resources with national standards and best practices. The new, shorter phone number makes it easier for people to remember and access mental health crisis services (though the previous 1 (800) 273-8255 number will continue to function indefinitely) and represents a continued commitment to those in crisis.

*Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-9355(WELL) – 24/7 on call*

Free and confidential therapy services, workshops, and training for situations related to gender- and power-based harm (including sexual assault, intimate partner violence, and stalking).

*Office for Equity, Equal Opportunity, and Title IX (EEO-TIX) - (213) 740-5086*

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

*Reporting Incidents of Bias or Harassment - (213) 740-5086 or (213) 821-8298*

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office for Equity, Equal Opportunity, and Title for appropriate investigation, supportive measures, and response.

*The Office of Student Accessibility Services (OSAS) - (213) 740-0776*

OSAS ensures equal access for students with disabilities through providing academic accommodations and auxiliary aids in accordance with federal laws and university policy.

*USC Campus Support and Intervention - (213) 740-0411*

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

*Diversity, Equity and Inclusion - (213) 740-2101*

Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

*USC Emergency* - *UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call*

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

*USC Department of Public Safety* - *UPC: (213) 740-6000, HSC: (323) 442-1200 – 24/7 on call*

Non-emergency assistance or information.

*Office of the Ombuds* - *(213) 821-9556 (UPC) / (323-442-0382 (HSC)*

A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

*Occupational Therapy Faculty Practice* - *(323) 442-2850 or* otfp@med.usc.edu

Confidential Lifestyle Redesign services for USC students to support health promoting habits and routines that enhance quality of life and academic performance.

### Resources for Online Students

The Course D2L page and the SSI Student Hub on D2L have many resources available for distance students enrolled in our graduate programs. In addition, all registered students can access electronic library resources through the link https://libraries.usc.edu/. Also, the USC Libraries have many important resources available for distance students through the link: https://libraries.usc.edu/faculty-students/distance-learners. These include instructional videos, remote access to university resources, and other key contact information for distance students.