

## **SSCI 575, Spatial Data Science**

### *Syllabus*

**Units:** 4

**Term Day Time:** Fall 2021, Wed & Fri, 11:00-12:50pm

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

**Instructor:** Orhun Aydin, PhD

**Office:** AHF B56G

**Regular Office Hours:** Wed 12-1 pm. Also available most days and times by appointment via email.

**Contact Info:** [oaydin@usc.edu](mailto:oaydin@usc.edu), (213) 740-5910, see contact page on Blackboard for Zoom Room

**Library Help:** Andy Rutkowski

**Office:** VKC 36B

**Office Hours:** Thu 10 a.m.-12 p.m.

**Contact Info:** [arutkows@usc.edu](mailto:arutkows@usc.edu), see contact page on Blackboard for Zoom Room

**IT Help:** Richard Tsung

**Office:** AHF 145D

**Office Hours:** By appointment

**Contact Info:** [spatial\\_support@usc.edu](mailto:spatial_support@usc.edu), 213-821-4415

## **Course Scope and Purpose**

Could we have prepared for an earthquake evacuation using data geotagged tweets? How are our movement patterns relate 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 that pertain to spatial machine learning methods frequently used in spatial analysis. Students will be 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. R and Python will be frequently utilized in class examples and projects.

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) 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 into different aspects of spatial machine learning and making connections with research and foundational ideas covered in class.

## ***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 multidisiplinary 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

## **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)

## **Course Structure**

The main theoretical concepts are introduced in class via lectures, discussions and after class via reading assignments.

Practical exercises utilize R, Python, and the ArcGIS platform for problem-solving. Assignments require coding in both languages, and working knowledge of R and 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**

ArcGIS 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/>

*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](mailto:spatial_support@usc.edu), making sure to copy (cc) me on the email.

*Communications* – All materials will be provided via Blackboard. This allows you to engage in reading and class preparation assignments both ahead of synchronous sessions or asynchronously. All assignments will be submitted via Blackboard. In addition to email about time-sensitive topics, announcements will be posted on the Blackboard Announcement page. It is each student's responsibility to stay informed as to course activities and updates. All students are in charge of ensuring that email sent from the USC Blackboard account is not directed to junk mail.

The instructor is regularly on-line and will respond to emails from students very quickly. 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 when the instructor is off-line for more than 72 hours, he will post an announcement on the Blackboard site.

*Discussion forums* – This course uses Piazza for the main means of discussion. 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 publically in the classroom. The instructor monitors the discussion threads and offer comments when necessary. Students should consider the discussion board a key way to connect with your classmates and share your discoveries.

## Required Readings and Supplementary Materials

The required 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 Blackboard. Due dates are shown in the summary that follows.

Resume Assignments – 2 worth a total of 5 points. All students are required to post and maintain a public resume, biography, and headshot on the SSI Student Community Blackboard site. Unless a student opts out, their resume will be included in the Spatial Sciences Institute Graduate Programs Resume Book, which is used to promote the program and highlight student skills, experiences, and professional aspirations. An updated resume will be submitted at the end of the course of study

Projects – 4 worth a total of 40 points. The projects 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 and Research Discussions – 5 worth a total of 25 points. These assignments call on students to critically analyze required readings, and discuss them with the instructor and their classmates during synchronous meetings and/or online discussion forums via Blackboard. Students are expected to answer critical thinking questions that provide students with an opportunity to apply their competencies to exploratory, open-ended scenarios.

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 20 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.

## Grading Breakdown

Assessment	Number	Points Each	Total Points
Weekly Assignments			
Resume Assignment	2	2.5	5
Projects	4	10	40
Reading Assignments	5	5	25
Mid-Term Exam	1	10	10
Final Exam	1	20	20
<b>Total</b>	<b>13</b>	<b>--</b>	<b>100</b>

## Assignment Submission Policy

Unless otherwise noted, assignments must be submitted via Blackboard by the due dates specified in the Course Schedule below and on the assignment instructions.

Unless otherwise noted, all assignments are *due by 11:59 pm Pacific Time (PT) on Mondays*. Project components have different due dates as indicated on 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.

## Schedule

	Topic	Readings and Assignments	Deliverables/Due Dates
<b>Week 1</b> 8/23	Is Spatial Special? Introduction to Spatial Data Science and GeoAI	Resume Assignment 1	No deliverables
<b>Week 2</b> 8/30	Introduction to the Spatial Data Science Workflow	Reading & Research Discussion 1 (RRD 1)	Resume Assignment 1 Due: Friday – 8/30
<b>Week 3</b> 9/6* <b>*Monday, 9/2 is university holiday</b>	Spatial Data Representation	Project 1	
<b>Week 4</b> 9/13	Spatial Data Wrangling		RRD 1 Due: Friday – 9/13 RRD 1 (synchronously, in class; asynchronously, before and after class session)

	<b>Topic</b>	<b>Readings and Assignments</b>	<b>Deliverables/Due Dates</b>
<b>Week 5</b> 9/20	Exploratory Spatial Data Analysis		
<b>Week 6</b> 9/27	Spatially Explicit Unsupervised Learning	Reading & Research Discussion 2 (RRD 2)	Project 1 Due: Friday – 9/27
<b>Week 7</b> 10/4	Spatially Explicit Supervised Learning (Linear Models)	Project 2	
<b>Week 8</b> 10/11* *10/14-10/15 is a university holiday	Spatially Explicit Neural Networks I	Reading & Research Discussion 3 (RRD 3)	RRD 2 Due: Friday – 10/11 RRD 2 (synchronously, in class; asynchronously, before and after class session)
<b>Week 9</b> 10/18	Spatially Explicit Neural Networks II		
<b>Week 10</b> 10/25	Model Assessment and Selection	Project 3	Mid-Term Exam (asynchronous)
<b>Week 11</b> 11/1	Spatial Optimization	Reading & Research Discussion 4 (RRD 4)	RRD 3 Due: Friday – 11/1 RRD 3 (synchronously, in class; asynchronously, before and after class session)
<b>Week 12</b> 11/8	Prescriptive Spatial Machine Learning	Reading & Research Discussion 5 (RRD 5)	Project 2 Due: Friday – 11/8
<b>Week 13</b> 11/15	Data Mining for Emerging Data Sources		
<b>Week 14</b> 11/22* *11/24-11/28 is a university holiday	Sentiment Analysis in Spatial Analysis	Project 4	Project 3, RRD 4 Due: Friday – 11/22 RRD 4 (synchronously, in class; asynchronously, before and after class session)
<b>Week 15</b> 11/29	Course Review	Resume Assignment 2	RRD 5 Due: Friday - 11/29 RRD 5 (synchronously, in class; asynchronously, before and after class session)

## Statement on Academic Conduct and Support Systems

### *Academic Conduct*

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. Please familiarize

yourself with the discussion of plagiarism in SCampus in Part B, Section 11, “Behavior Violating University Standards” <https://policy.usc.edu/files/2020/07/SCampus-Part-B-1.pdf>. Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on scientific misconduct, [policy.usc.edu/scientific-misconduct](https://policy.usc.edu/scientific-misconduct).

### **Support Systems**

*Counseling and Mental Health* – (213) 740-9355 – 24/7 on call  
[engemannshc.usc.edu/counseling](https://engemannshc.usc.edu/counseling)

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

*National Suicide Prevention Lifeline* – 1 (800) 273-8255 – 24/7 on call  
[www.suicidepreventionlifeline.org](https://www.suicidepreventionlifeline.org)

Free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week.

*Relationship and Sexual Violence Prevention Services (RSVP)* – (213) 740-9355(WELL), press “0” after hours – 24/7 on call  
[studenthealth.usc.edu/sexual-assault](https://studenthealth.usc.edu/sexual-assault)

Free and confidential therapy services, workshops, and training for situations related to gender-based harm.

*Office of Equity and Diversity (OED)* – (213) 740-5086 | *Title IX Compliance* – (213) 821-8298  
[equity.usc.edu](https://equity.usc.edu), [titleix.usc.edu](https://titleix.usc.edu)

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  
[usc-advocate.symlicity.com/care\\_report](https://usc-advocate.symlicity.com/care_report)

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

*The Office of Disability Services and Programs* – (213) 740-0776  
[dsp.usc.edu](https://dsp.usc.edu)

Support and accommodations for students with disabilities. Services include assistance in providing readers/notetakers/interpreters, special accommodations for test taking needs, assistance with architectural barriers, assistive technology, and support for individual needs.

*USC Campus Support and Intervention* – (213) 821-4710  
[campussupport.usc.edu](https://campussupport.usc.edu)

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



*Diversity at USC – (213) 740-2101*

[diversity.usc.edu](https://diversity.usc.edu)

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*

[dps.usc.edu](https://dps.usc.edu), [emergency.usc.edu](https://emergency.usc.edu)

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) 442-120 – 24/7 on call*

[dps.usc.edu](https://dps.usc.edu)

Non-emergency assistance or information.

### ***Resources for Online Students***

The Course Blackboard page and the GIST Community Blackboard page 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.