# CE 263N **Scalable Spatial Analytics**

Lecture: MW 11 – 12 pm, 212 O'Brien
Seminar: F 9 – 10 am, 212 O'Brien

### Prerequisites

No formal prerequisites. (an undergraduate degree in engineering domain is expected, with knowledge of linear algebra and statistics, as well as some programming experience)

**Primary goal:** Provide theoretical background and hands-on experience in predictive modelling for designing scalable data-driven systems and location-based services based on data analytics.

**Short Description:** Introduction to modern methods of data analysis, spatial data handling and visualization technologies for engineers and data scientists. Theoretical coverage includes a selection of methods from spatial statistics, exploratory data analysis, spatial data mining, discriminative and generative approaches of machine learning. Projects and assignment tasks are targeted at real-world scalable implementation of systems and services based on data analytics in environmental remote sensing, transportation, energy, location-based services and the domain of "smart cities" in general.

**Instructor**: Prof Alexey Pozdnukhov
Email: alexeip@berkeley.edu
Phone: (510) 642-1008
Office: 115 McLaughlin Hall
OH: M W F 4:00-5:30

### Course Reader

The course readings are available on bCourses as pdf files and/or links to online materials.

### Course Grading

| | |
|---|---|
| 1 or 2 Midterms | 30% |
| Homework (6 problems) | 50% |
| Final Project | 20% |

## Course Schedule

| LECTURE | DATE | TOPIC | THEME | PROBLEMS DUE |
|---|---|---|---|---|
| 1 | 9/03 | Introduction | General | |
| 2 | 9/08 | Exploratory data analysis | EDA | |
| 3 | 9/10 | Clustering | EDA | |
| 4 | 9/15 | Dimensionality reduction | EDA | |
| 5 | 9/17 | Spatial databases I | Data Handling | |
| 6 | 9/22 | Spatial databases II | Data Handling | |
| 7 | 9/24 | GIS and web mapping | Data Handling | |
| 8 | 9/29 | Trajectories and map matching | Data Processing | |
| 9 | 10/01 | Map matching and reconstruction | Data Processing | |
| 10 | 10/06 | Point Pattern Analysis | Spatial Statistics | |
| 11 | 10/08 | Geostatistics | Spatial Statistics | |
| 12 | 10/13 | Spatial Regression | Spatial Statistics | |
| 13 | 10/15 | Geographically weighted models | Spatial Statistics | |
| 14 | 10/20 | Intro to ML, model selection | ML discriminative | |
| 15 | 10/22 | Locally linear models | ML discriminative | |
| 16 | 10/27 | Kernel methods I, SVM | ML discriminative | |
| 17 | 10/29 | Kernel methods II, SVR | ML discriminative | |
| 18 | 11/03 | Elements of computer vision | ML discriminative | |
| 19 | 11/05 | Mixture models, EM | ML generative | |
| 20 | 11/10 | MC, HMM, CRF | ML generative | |
| 21 | 11/12 | Spatial PLSA/LDA | ML generative | |
| 22 | 11/17 | Recommender Systems | ML generative | |
| 23 | 11/19 | Complex networks | Complex Systems | |
| 24 | 11/24 | Agent-based models | Complex Systems | |
| 25 | 11/26 | Catch-up | General | |
| 26 | 12/01 | Final projects | General | |
| 27 | 12/03 | Final projects | General | |
| | 5/5 | Review (Reading Week) | | |
| | 5/7 | Review (Reading Week) | | |

**Tentative Reading Schedule**

| SESSION | DATE | READINGS |
|---|---|---|
| 1 | 09/05 | Breiman L., Statistical Modeling: The Two Cultures, Statistical Science, Vol. 16, No. 3, pp. 199–23, 2001. |
| 2 | 09/08 | EDA online resources |
| 3 | 09/10 | Scully D., Web-Scale K-Means Clustering. WWW'2010.<br>Bottou L., Bengio Y., Convergence Properties of the k-Means Algorithms, NIPS'95.<br>Ester M, Kriegel H-P., Sander J., Xu X., A density-based algorithm for discovering clusters in large spatial databases with noise. KDD-96 |
| 4 | 09/15 | Burges C., Dimension Reduction: A Guided Tour, 2009 (selected chapters) |
| 5 | 09/17 | Web resources: Postgre/PostGIS |
| 6 | 09/22 | Web resources: QGIS |
| 7 | 09/24 | Web resources: MongoDB |
| 8 | | |
| 9 | | |
| 10 | | |

**Seminar Schedule**

09/05 Statistical modelling: the two cultures. Assignments problems overview.

09/12 Problem 1 introduction: clustering

09/19 Clustering and dimensionality reduction examples

09/26 Problem 2 introduction: PostgreSQL / PostGIS, MongoDB

10/03

10/10 Problem 3 introduction

10/17

10/24 Problem 4 introduction

10/31

11/07 Problem 5 introduction

11/14

11/21 Problem 6 introduction

11/28

12/01 Final project presentations
12/03 Final project presentations