

UCSB Data Science Bootcamp 2015

A two-week course, held just before the start of the academic year, meant to introduce and refresh skills around programming, software, and data.

Supported by the Network Science IGERT through the National Science Foundation (Grant# DGE-1258507), and by the College of Engineering and the College of Letters and Science at the University of California, Santa Barbara.

Registration:

<http://goo.gl/forms/1ikPkQTSaI>

Location:

Engineering Sciences Building, Cooper Lab.

Schedule:

Mon 9/7	Tues 9/8	Weds 9/9	Thurs 9/10	Fri 9/11
Labor Day	9:30 am Program Overview 11:30 am Assessment 1:30 pm Computer Basics	9:00 am Everything Data 12:00 pm Faculty Lunch 1:30 pm Small Data, Big Data	9:00 am Visualization I 1:30 pm Visualization II	9:00 am Linear Algebra

Mon 9/14	Tues 9/15	Weds 9/16	Thurs 9/17	Fri 9/18
9:00 am Intro to Algorithms 12:00 pm Faculty Lunch 1:30 pm Graph Algorithms I	9:00 am Graph Algorithms II 1:30 pm Dynamical Systems	9:00 am Foundations of Statistics 12:00 pm Faculty Lunch 1:30 pm Intro to Machine Learning	9:00 am Machine Learning on Graphs 4:00 pm Hands-on Machine Learning (Note time change)	9:00 am Graduate Student Panel

Course Material can be found on Luca Foschini's Git Repository:

<https://github.com/LucaFoschini/UCSBDataScienceBootcamp2015>

Introduction to Bootcamp

Date: 9/8/15

Time: 9:30 am -- 11:30 am

Format: Panel

Instructors: Luca Foschini, John Mohr, Steven Proulx, Ambuj Singh

Module Goals/Outline:

Panel discussion and overview of the boot camp. Each of the panelists will discuss his research and express views on the need for Network Science. The presentations will be informal and informative. There will be no need for power point or other materials. Luca Foschini will then present the bootcamp structure and provide an overview of the various modules.

The Introduction will be followed by an informal assessment of the boot camp participants at 11:30 am.

Computer Basics

Date: 9/8/15

Time: 1:30 pm -- 3:30 pm

Instructor: Haraldur Tómas Hallgrímsson

Format: Hands-on Lab

Module Goals/Outline:

- Unix Basics
 - How to open and use the terminal
 - How to connect to Unix servers (ssh)
 - Text manipulation and command-line magic
- Git
 - The importance of version control
 - Github, reproducibility and the scientific method
- Python and Jupyter notebooks
 - Introduction to Python
 - Jupyter (née IPython) notebooks
 - Will be used throughout rest of the bootcamp

Everything Data

Date: 9/9/15

Time: 9:00 am -- 12:00 pm

Instructor: Luca Foschini

Format: Hands-on, Lab

Module Goals/Outline:

- What is data? Data representation in a computer
- Native data types in Python: integer, list, dict, numpy arrays, pandas
- From simple to complex: text, time series, networks, geometric objects
- Discuss complexity of manipulation of these objects.
- Load and visualize different datasets on Python

Lunch with Faculty

Date: 9/9/15

Time: 12 pm

Instructor: Prof. Francesco Bullo

Format: Catered Lunch

Module Goals/Outline:

Prof. Bullo will discuss his research and views of network science over an informal lunch.

Small Data, Big Data

Date: 9/9/15

Time: 1:30 pm -- 3:30 pm

Instructor: Luca Foschini

Format: Hands-on, Lab

Module Goals/Outline:

- Small data, big data. Do you really have big data?
- Storage latencies. cache/ram/SSD/Redis/s3
- Computation engines: single core, multi-core, memory distributed, disk distributed. Pandas, numpy multi-core extensions
- Examples: single-machine SSD-backed operation. Caveats (sequential access needed)

Visualization I	
<p>Date: 9/10/15 Time: 9:00 am -- 12:00 pm Instructor: John O'Donovan Format: Hands-on, Lab</p>	<p>Module Goals/Outline:</p> <ul style="list-style-type: none">● Introduction to Visualization -visual variables, design, types etc.● Visualization in Academia and Industry● Visualizing Live Twitter Data (Demo)● Anatomy of a research paper in Visualization

Visualization II	
<p>Date: 9/10/15 Time: 1:30 pm -- 3:30 pm Instructor: John O'Donovan Format: Hands-on, Lab</p>	<p>Module Goals/Outline:</p> <ul style="list-style-type: none">● Visualizing live feeds using Python and Plot.ly (Hands-on)● Visualizing network data (Fluo, TopicNets Demos)● Visualizing data in R (Hands on)● Interactive Visualization (Demo)

Linear Algebra	
<p>Date: 9/11/15 Time: 9:00 am -- 11:00 am Instructor: Victor Amelkin Format: Hands-on, Lab</p>	<p>Module Goals/Outline:</p> <ul style="list-style-type: none">● Review of Linear Algebra's Fundamentals<ul style="list-style-type: none">○ Matrix algebra○ Vector spaces○ Angles, lengths, projection○ Eigenproblem, SVD● Linear Algebra and Graphs<ul style="list-style-type: none">○ Graphs: definitions, properties, representation○ Graph algorithms in the language of linear algebra○ Spectral graph theory

Introduction to Algorithms

Date: 9/14/15

Time: 9:00 am -- 12:00 pm

Instructor: Luca Foschini

Format: Lecture

Module Goals/Outline:

- Basic data structures: arrays, lists, balanced binary trees (sets), hash tables (dicts);
- some organized/categorical view of what each of these data structures are good for, and how they compare in cost/benefit
- NP Hardness (TSP vs Eulerian Path)

Lunch with Faculty

Date: 9/14/15

Time: 12 pm -- 1 pm

Instructor: Prof. Alberto Busetto

Format: Brown bag lunch

Module Goals/Outline:

Prof. Busetto will discuss his research and views of network science over an informal lunch.

Graph Algorithms I

Date: 9/14/15

Time: 1:30 am -- 3:30 pm

Instructor: Luca Foschini

Format: Hands-on

Module Goals/Outline:

- Graph definitions (directed, undirected, weighted, unweighted, trees, cycles, bipartite, complete) etc...
- Graph representations (adjacency matrix, adjacency list); pros and cons;
- Graph generation: Erdos-Renyi model
- Simple graph definition in Python NetworkX

Graph Algorithms II

Date: 9/15/15

Time: 9:00 am -- 12:00 pm

Instructor: Nirman Kumar

Format: Hands-on

Module Goals/Outline:

- Algorithms on Graphs.
- Introduction to special classes of graphs
- Demonstration of a few algorithms above in networkX
- Examples in NetworkX on real networks (social, brain)
- Preferential attachment
- Small-world networks
- Hands-on
 - Generating and characterizing several graphs (both synthetic and real)
 - Counting triangles
 - Visualization using GraphViz

Dynamical Systems

Date: 9/15/15

Time: 1:30 am -- 3:30 pm

Instructor: Hari Sivakumar

Format: Hands-on

Module Goals/Outline:

- Introduction to dynamics
- Why is it important to study dynamics on networks?
- First order dynamics - flows on a line
- Linear vs nonlinear dynamics
- Stability analysis -an intuitive explanation
- Lyapunov equations - an intuitive explanation
- Introduction to second order dynamics
- Eigenvalues and eigenvectors, stability
- Introduction to bifurcations and hysteresis

Foundations of Statistics

Date: 9/16/15

Time: 9:00 am -- 12:00 pm

Instructor: Arya Pourzanjani

Format: Lecture

Module Goals/Outline:

- Basic probability and combinatorics.
- Bernoulli trials. Expectation. Variance. Tail bounds.
- Significance and p-values
- Regression, controlling, example in R

Lunch with Faculty

Date: 9/16/15

Time: 12 pm -- 1 pm

Instructor: Prof. Susan Cassels

Format: Brown bag lunch

Module Goals/Outline:

Prof. Cassels will discuss her research and views of network science over an informal lunch.

Introduction to Machine Learning

Date: 9/16/15

Time: 1:30 am -- 3:30 pm

Instructor: Bo Zong

Format: Hands-on

Module Goals/Outline:

- Supervised learning
 - Decision Tree and Random Forest
 - Linear Regression and Support Vector Machine
 - Logistic Regression and Neural Network
- Unsupervised learning
 - k-Means, k-Medoids, and Hierarchical Clustering
 - Mixture Modeling

Machine Learning on Graphs

Date: 9/17/15

Time: 9:00 am -- 12:00 pm

Instructor: Bo Zong

Format: Hands-on

Module Goals/Outline:

- Classification on graphs
- Community detection
- Frequent patterns

Hands-on Machine Learning

Date: 9/17/15

Time: 1:30 am -- 3:30 pm

Instructor: Luca Foschini

Format: Hands-on

Module Goals/Outline:

- Fun with scikit-learn.
- Deep learning example

Graduate Student Panel

Date: 9/18/15

Time: 9:00 am

Instructors: Alyssa Newman, Hari Sivakumar, Herbert Cai, Bo Zong

Format: Panel

Module Goals/Outline:

This panel will give the boot camp participants a chance to hear graduate students talk about their research and the value of a research background in network science.