



# School of Professional Studies

## **DATA 622 Big Data and Machine Learning - Fall 2019**

Credits: 3 cr. Prerequisite(s): DATA 621

### **Summary:**

This course explores big data concepts and machine learning models widely used in various industries and introduces students to the fundamental concepts of big data engineering.

At the end of this course, students will be able to:

- Apply and evaluate advanced modeling techniques such as ensemble modeling, recommendation engine, neural networks, and deep learning.
- Understand and implement basic data engineering concepts such as distributed computing, processing, model deployment, and data pipelines.

### **Lectures:**

This course will be 100% online. Lectures will be conducted live via gotomeeting. There will be 7 lectures, scheduled bi-weekly on Wednesday evenings.

### **Contact Information:**

Professor Haiyuan Wang

You may reach me via email at [haiyuan.wang@sps.cuny.edu](mailto:haiyuan.wang@sps.cuny.edu) for grades and other academic concerns.

For homework and course content related questions, I highly recommend you use our class's Slack chat to ask me questions so that your classmates could also benefit from the information. You will be given setup information for Slack in class.

### **Grading Policy:**

4 homework 70%

1 project 30%

Homework and project will be a mix of required reading, recommended reading, coding exercises, and concept questions.

### **Late Policy:**

Late homework and project submissions are due at 11:59 PM EST on the due date specified. Please reach out early in case you need late submission with proper reason.

There will be a 10% penalty for each day that your submission is late.

### **Required Reading:**

Deep Learning Book (free): <http://www.deeplearningbook.org/>

## Schedule:

Unit	Topic	Readings	"To Do"
<b>Unit 1</b>			
Week 1 - 2 Aug 27-Sep 8	<b>Topics:</b> Basic tech stack: bash, Python, git, SQL <b>Lecture:</b> 8/28	Please see lecture notes for reading instructions	
Week 3-4 Sep 9-22	<b>Topics:</b> Data pipeline, data processing <b>Lecture:</b> 9/11	Please see lecture notes for reading instructions	homework 1 is due by EOW (Sunday 11:59pm)
<b>Unit 2</b>			
Week 5-6 Sep 23-Oct 6	<b>Topics:</b> Model deployment, agile data science <b>Lecture :</b> 9/25	Please see lecture notes for reading instructions	homework 2 is due by EOW (Sunday 11:59pm)
Week 7-8 Oct 7-20	<b>Topics:</b> Data infrastructure <b>Lecture :</b> 10/09	Please see lecture notes for reading instructions	homework 3 is due by EOW (Sunday 11:59pm)
<b>Unit 3</b>			
Week 9-10 Oct 21- Nov 3	<b>Topics:</b> Dimension reduction & linear optimization <b>Lecture :</b> 10/23	Please see lecture notes for reading instructions	
Week 11-12 Nov 4-17	<b>Topics:</b> Distributed file systems <b>Lecture :</b> 11/6	Please see lecture notes for reading instructions	homework 4 is due by EOW (Sunday 11:59pm)
Week 13-14 Nov 18-Dec 1	<b>Topics:</b> Ensemble modeling <b>Lecture :</b> 11/20	Please see lecture notes for reading instructions	
Week 15-16 Dec 2–Dec 15	<b>Topics:</b> Neural networks & deep learning <b>Lecture :</b> 12/4	Please see lecture notes for reading instructions	Final project report due by end of week 16th

## ACCESSIBILITY AND ACCOMMODATIONS

The CUNY School of Professional Studies is firmly committed to making higher education accessible to students with disabilities by removing architectural barriers and providing programs and support services necessary for them to benefit from the instruction and resources of the University. Early planning is essential for many of the resources and accommodations provided. Please see:

[http://sps.cuny.edu/student\\_services/disabilityservices.html](http://sps.cuny.edu/student_services/disabilityservices.html)

## ONLINE ETIQUETTE AND ANTI-HARASSMENT POLICY

The University strictly prohibits the use of University online resources or facilities, including Blackboard, for the purpose of harassment of any individual or for the posting of any material that is scandalous, libelous, offensive or otherwise against the University's policies. Please see:

[http://media.sps.cuny.edu/filestore/8/4/9\\_d018dae29d76f89/849\\_3c7d075b32c268e.pdf](http://media.sps.cuny.edu/filestore/8/4/9_d018dae29d76f89/849_3c7d075b32c268e.pdf)

## ACADEMIC INTEGRITY

Academic dishonesty is unacceptable and will not be tolerated. Cheating, forgery, plagiarism and collusion in dishonest acts undermine the educational mission of the City University of New York and the students' personal and intellectual growth. Please see: [http://media.sps.cuny.edu/filestore/8/3/9\\_dea303d5822ab91/839\\_1753cee9c9d90e9.pdf](http://media.sps.cuny.edu/filestore/8/3/9_dea303d5822ab91/839_1753cee9c9d90e9.pdf)

## STUDENT SUPPORT SERVICES

If you need any additional help, please visit Student Support Services:

[http://sps.cuny.edu/student\\_resources/](http://sps.cuny.edu/student_resources/)