

Course Syllabus

CS5661: **Topics in Data Science**, Prerequisite: CS4661.
Advanced algorithms, techniques, and tools for machine learning, data processing, and extracting knowledge from large-scale data, and use it for future purposes such as prediction, detection, and decision making.
This course will particularly focus on Artificial Neural Networks, Deep Learning, Convolution Neural Networks, and other advanced machine learning methods, as well as techniques in data processing, data cleaning/scraping, data analytics, dimensionality reduction, and visualization.

Class Time and Location: Friday, 11:30AM – 2:00PM, ASCB 132

Instructor: Mohammad Pourhomayoun
Email: mpourho@calstatela.edu
Office: ET A-408
Office Hours: Wed 1:30PM – 2:30PM, Fri 10:00AM-11:30AM.

Textbook: There is no required textbook. Nonetheless, we will introduce many optional references in the class for interested students.

Student learning outcomes: Students who successfully complete this course will be able to:

- Understand and use advanced techniques for *data pre-processing*.
- Understand advanced concepts, applications, and algorithms of *data analytics, data mining, and machine learning*.
- Understand and use the most important statistical modeling and machine learning algorithms including *Support Vector Machine (SVM), Artificial Neural Networks (ANN), Deep Neural Networks, Principal Component Analysis, Convolutional Neural Networks and Deep Learning*, etc.
- Understand the advantages/disadvantages of the new techniques in comparison with previously learned methods
- Understand and use various algorithms for *dimensionality reduction and feature selection for big data analytics*.
- Understand and use *result visualization/validation, error identification, cross validation, various metrics for evaluation the prediction accuracy*.
- Understand the basics of *big data manipulation, parallel processing, MapReduce*, and main tools and frameworks.
- Understand and use the main algorithms and techniques for *large-scale data analytics and large-scale machine learning*.
- Understand and use tools, software packages, and libraries for data processing, analytics, and machine learning.

Grading:

- Assignments: 40%
 - Theoretical Problems
 - Implementation and Programming
 - Assignments are due at the beginning of class on the due date.
 - Late submissions will not be accepted.
- Final Project: 20%
- Final Exam: 40%
- Participation: 5%

ADA statement: Reasonable accommodation will be provided to any student who is registered with the Office of Students with Disabilities and requests needed accommodation.

Academic honesty statement: Students are expected to do their own work and to abide by the University Policy on academic honesty, which is stated in the Schedule of Classes. Copying the work of others, cheating on exams, and similar violations will be reported to the University Discipline Officer, who has the authority to take disciplinary actions against students who violate the standards of academic honesty.

Student responsibilities: Students are responsible for being aware of all announcements that are made in class, such as changes in due dates of homework and papers, and cancellation of class due to instructor's absence. Students are responsible for announcements made on days that they are absent.

Students must check their CSULA email account regularly for information from the instructor and the Department. Failure to do so may result in missed deadlines or other consequences that might adversely affect students. Note that you can forward this email account to any other account of your choosing.