MASC 499

Instructor: Priya Vashishta

Introduction to Machine Learning for Materials

Spring 2021, 4 Units, M&W, 4:00-5:50pm

- ✓ No prerequisite – Simple vectors & matrices and 1st & 2nd derivative and Gaussian Integral are used
- ✓ Simple homework & three tests on course material
- ✓ Hands-on projects on machine learning
- ✓ No midterm Exam; No Final Exam

In this course, you will learn,

- Machine learning concepts and models
- Data science tools and Python programming
- How to apply ML technique to solve real-world material-related problems

Contact Anthony Tritto at artritto@usc.edu for d-clearance
https://classes.usc.edu/term-20211/classes/masc

Course Description

This is an introductory level course on machine learning (ML) for materials applications. The rapid evolution and exponential growth of machine learning in materials science and engineering is evident in novel data science skillset and machine learning knowledge is necessary to be competitive in the future materials science and engineering fields. The goal of the course is to provide engineering students basic concepts of machine learning, statistics and probability, and teach them modelling techniques for materials and processes to solve real-world engineering problems.

This is an introductory course to introduce ML concepts and techniques to engineering students with basic engineering backgrounds. The course will focus on the applications of ML methods and its application to materials and less on the intricate mathematical details.

Learning Objectives

Students will learn basic knowledge of machine learning and its application to materials and gain hands-on experience in order to:

1. Understand machine learning concepts, models, their applicability.
2. Learn how to develop a program using Python and data science tools.
3. Come up with engineering real-world problem that involve materials and apply ML technique to solve the problem.

Prerequisite(s): None
Co-Requisite(s): None
Concurrent Enrollment: None
Recommended Preparation: Basic Python programming and basic calculus.

Course Notes

Grading type: letter grade. All course notes will be provided on Blackboard.

Technological Proficiency and Hardware/Software Required

Personal laptop computer is necessary to access the USC's Center for Advanced Research Computing (CARC) for the class projects. Students will be provided simulation programs and individual CARC account will be setup for all students in the class.

Required Readings and Supplementary Materials

An Introduction to Statistical Learning (ISL), G. James, D. Witten, and T. Hastie. ISBN-13: 978-1461471370

Lecture notes and other reading materials will be given in the class and also posted on Blackboard.

Description and Assessment of Assignments

The learning outcome is assessed by homework, midterm exam, and class project.

Homework: Bi-weekly homework assignments including Python programming are due at the beginning of class. Late submission is penalized by 25% and will not be accepted after 24 hours past the due date and time. The lowest homework score will be dropped from the final grade.

Exam and Project: One midterm exam and class projects

Project Objective:

The purpose of the class project is to understand basics of materials modeling using ML techniques and apply it to real-world problems. Each student needs to come up with a project idea and discuss with instructor during office hours. Discussion with other students is encouraged, however, every student must complete project implementation, analyze and visualize simulation results, and submit own project report. During the final week, each student will meet with the instructor to review the project reports. The final project review is mandatory and partly contributes toward the project grade.

Grading Breakdown

Assignment	Points	% of Grade
Homework	100	40
Midterm	100	20
Final Project	100	40
Total	100	100

Assignment Submission Policy

Final project report and homework must be submitted by the due date announced in class.

Course Breakdown

- **Basic Math:** Review of Linear Algebra, Statistics, and Probability
- **Programming and Data Science Tool:** Introduction to Python, Materials Databases
- **Linear Regression:** Univariate, Multivariate, Polynomial Regressions
- **Classification:** Logistic Regression, Cross-validation, Ridge and Lasso Regressions
- **Unsupervised Learning**: K-means and PCA
- **Three-Based Methods:** Regression and Classification Trees
- **Support Vector Machine:** Maximum margin classifier, Soft Margin, Kernel method
- **Theory of Learning:** Bounding of Growth Dimension, VC dimension
- **Deep Neural Network:** Multilayer perceptron, Convolutional and Recurrent Neural Network