College of Science · Computer Science

Introduction to Machine Learning Section 04

CS 171

Fall 2023 3 Unit(s) 08/21/2023 to 12/06/2023 Modified 08/18/2023

Contact Information

Instructor	Dr. Tahereh Arabghalizi
Office	MH 217
Email	tahereh.arabghalizi@sjsu.edu
Office Hours	M/W 12-1:15 pm and TU/TH 3-4:15 pm (Email (mailto:tahereh.arabghalizi@sjsu.edu) OR by appointment (https://calendly.com/tahereh-arabghalizi-sjsu-fall23)) - All questions should be asked during the office hours, unless they are short/simple questions.
Teaching Assistant	TBD

Course Description and Requisites

Covers a selection of classic machine learning techniques including backpropagation and several currently popular neural networking and deep learning architectures. Hands-on lab exercises are a significant part of the course. A major project is required.

Prerequisite(s): CS 146 (with a grade of C- or better). Computer Science or Software Engineering majors only.

Letter Graded

Classroom Protocols

Students are expected to adhere to the Student Conduct Code found at the SJSU Student Conduct website (http://www.sjsu.edu/studentconduct/) (http://www.sjsu.edu/studentconduct/). Additionally, students should regularly attend lectures and labs (if applicable), treat instructors and peers with respect, and refrain from the use of cell phones during any classroom activities.

- Regular class attendance is highly recommended and strongly encouraged.
- Please arrive to class on time so that you benefit fully from the course experience and you do not disturb classmates and the instructor while class is in session.
- Students are responsible for knowing all materials covered in class lectures, readings, assignments, and other course-related work.
- Laptops, tablets, and other devices should only be used for course-related purposes.

Program Information

Diversity Statement - At SJSU, it is important to create a safe learning environment where we can explore, learn, and grow together. We strive to build a diverse, equitable, inclusive culture that values, encourages, and supports students from all backgrounds and experiences.

Course Goals

This course is an Introduction to the fundamental principles and practical applications of Machine Learning. You'll gain a good understanding of both supervised and unsupervised learning techniques, exploring algorithms like linear regression, decision trees, SVM, k-means clustering, Neural Networks, large language models and more. Throughout the course, you will learn effective strategies to overcome common challenges in machine learning such as overfitting, imbalanced data, etc. ensuring the robustness and reliability of your models.

Course Learning Outcomes (CLOs)

After completing this course students should have a working knowledge of a wide variety of machine learning topics and have a good understanding of how to apply such techniques to real-world problems.

Course Materials

Recommended Texts/Readings

- An introduction to statistical learning with applications in Python by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, Jonathan Taylor
- Pattern Recognition and Machine Learning by Christopher M. Bishop
- Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow by Aurélien Géron
- Python Crash Course by Eric Matthes

Other technology requirements / equipment / material

- You will need a wireless laptop with internet access.
- Python 3, Scikitlearn libraries, numpy/scipy, tensorflow/keras, Jupyter notebooks, Google Colab, etc.

Course Requirements and Assignments

- **Reading Assignments**: You may be assigned readings from the textbooks or papers/articles.
- **Videos**: Videos maybe be posted to introduce new topics not covered in class or as supplementary materials.
- **Unannounced in class exercises, pop quizzes/questions and discussion forum** may be given anytime. The purpose of in class exercises and pop questions is to encourage you to learn, study and review the concepts and materials presented/discussed in the lecture.
- **Programming Assignments**: Programming assignments are to be done individually, unless otherwise specified. They can be discussed but should be implemented individually. More information is given at the time of the first programming assignment. Never use any code you find on the web, unless I provide it. Some assignments may have an oral discussion or examination.
- **Final Project**: Students can either form a team with 2 members or individually develop ML/DL/LLM models. Students have the flexibility to either choose their own areas of interest or receive guidance from the lecturer. Apply the learned knowledge starting from data preprocessing to evaluation. Each team should submit the final project including three deliverables:
    1. A well-structured code with clear comments. The code should be able to be run through without syntax error, logic error, and not counter business sense.
    2. A report that provides Project Introduction/ Motivation, Problem Statement, Data Source and Preprocessing, Methodology, Model Training and Performance, Challenges Faced, Results and Interpretation, Future Directions, Conclusion, and References.
    3. A presentation with the same aspects mentioned in the report.
- **Midterm Exam**: there will be one exam (midterm) during the semester.
- **Final Exam**: The final exam could be comprehensive or non-comprehensive (TBD).

Course Policies

Incomplete Work: Points will be deducted for incomplete question responses and solutions that are partially functional. Consult individual assignments for details of point allocation for each problem.

Late Assignments: No late homework will be accepted. However, under exceptional circumstances, one problem set per student might be accepted late. It will need to be handed in before the following class meeting and will be graded with 30% off. Such an extension should be requested from the instructor.

Makeup Exams: Makeup exams will only be given in cases of illness (documented by a doctor) or in cases of documentable, extreme emergencies.

Academic Honesty: Students must only submit their own work for all quizzes, assignments, exams, and projects. Copying and any other form of cheating will not be tolerated and will result in a failing grade (F) for the course, as well as disciplinary consequences from the university.

Grading Information

Success in this course is based on the expectation that you will spend, for each unit of credit, a minimum of 45 hours over the length of the course (normally three hours per unit per week) for instruction, preparation, and studying. Plan on spending at least 7 hours per week outside of lecture time engaging with the course material.

Grading Information

Course weightings will be as follows:

- 25% Programming Assignments
- 25% Final Project
- 25% Midterm Exam
- 25% Final Exam

Your course grade will be determined by your final weighted average:

- *A plus = 97% or higher*
- *A = 93% up to 97%*
- *A minus = 90% to 93%*
- *B plus = 87% to 90%*
- *B = 83% to 87%*
- *B minus = 80% to 83%*
- *C plus = 77% to 80%*
- *C = 73% to 77%*
- *C minus = 70% to 73%*
- *D plus = 67% to 70%*
- *D = 63% to 67%*
- *D minus = 60% to 63%*
- *F = 0% to 60%*
- Boundary cases count as the higher of the two grades.

All students have the right, within a reasonable time, to know their academic scores, to review their grade-dependent work, and to be provided with explanations for the determination of their course grades. See [University Policy S20-2] for more details.

University Policies

Per [University Policy S16-9 (PDF) (http://www.sjsu.edu/senate/docs/S16-9.pdf)], relevant university policy concerning all courses, such as student responsibilities, academic integrity, accommodations, dropping and adding, consent for recording of class, etc. and available student services (e.g. learning assistance, counseling, and other resources) are listed on the [Syllabus Information (https://www.sjsu.edu/curriculum/courses/syllabus-info.php)] web page. Make sure to visit this page to review and be aware of these university policies and resources.

Course Schedule

The course schedule is subject to change with fair notice. Changes will be announced on Canvas.

Week	Topics
1	Course Introduction, Prerequisites Check
1	Introduction to Machine Learning
2	Supervised Learning & Decision Trees
2	Decision Trees (continued)
3	Overfitting
3	k-Nearest Neighbor & Evaluation
4	Linear Regression & Gradient Descent
4	Regularization
5	Logistic Regression
5	Learning Theory OR Dealing with Imbalanced Data
6	Support Vector Machines & Kernels
6	Ensemble Methods
7	Review for Midterm
7	**Midterm Exam**
8	Probability Review
8	Naive Bayes
9	Text Classification & Evaluation
9	Text Classification & Evaluation (continued)
10	Neural Networks
10	Deep Learning
11	Deep Learning (continued)

Week	Topics
11	Advanced Topics in Deep Learning (LLM)
12	Unsupervised Learning
12	Reinforcement Learning
13	Reinforcement Learning (continued)
13	Principal Components Analysis (PCA)
14	Review for Final Exam
14	No Class (Thanksgiving)
15	Project Presentations
15	Project Presentations
Final Exam	Check Date and Time Here (https://www.sjsu.edu/classes/final-exam-schedule/fall-2023.php)