SJSU SAN JOSÉ STATE UNIVERSITY

College of Science · Computer Science

Introduction to Machine Learning Section 03

CS 171

Fall 2024 3 Unit(s) 08/21/2024 to 12/09/2024 Modified 08/20/2024

Contact Information

Instructor: Saptarshi Sengupta, PhD

Office Location: MacQuarrie Hall 416

Telephone: 408-924-4808

Email: saptarshi.sengupta@sjsu.edu

Office Hours: Friday, 11:00 AM – 1:00 PM

Class Days/Time: TuTh 7:30PM - 8:45PM

Classroom: DH 450

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

- Cheating will not be tolerated.
- Student must be respectful of the instructor and other students. For example, No disruptive or annoying talking.
- Turn off cell phones
- Class begins on time

- Valid picture ID required at all times

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 Learning Outcomes (CLOs)

The focus of this course will be machine learning, with examples from fields such as Cyber-Physical Systems, Natural Language Processing, Computer Vision, Anomaly Detection, Health Informatics and Prognostics. After completing this course students will be able to implement a variety of supervised, unsupervised and reinforcement learning models in a problem of interest. Upon completion students will also be able to work on real-world problems relevant to the industry using the insights learned in the class.

Course Materials

Textbook

Machine Learning: An Algorithmic Perspective, Second Edition, 2014, Chapman and Hall/CRC

Authors: Stephen Marsland

ISBN-13 : 978-1-4665-8333-7 (eBook - PDF)

Other Readings

Machine Learning with Applications in Information Security, by Mark Stamp, published by Chapman Hall/CRC in 2017.

ISBN-10 : 1138626783, ISBN-13 : 978-1138626782

Deep Learning (Adaptive Computation and Machine Learning series)

Authors: Ian Goodfellow, Yoshua Bengio, Aaron Courville

ISBN-13: 9780262035613, ISBN-10: 0262035618

Course Requirements and Assignments

Course Format

Faculty Webpage and mySJSU messaging

Course materials such as syllabus, handouts, notes, assignment instructions, etc. can be found on Canvas Leaning Management System course login website at http://sjsu.instructure.com. You are responsible for regularly checking with the messaging system through MySJSU at http://my.sjsu.edu (or other communication system as indicated by the instructor) to learn of any updates.

Course Requirements and Assignments

SJSU classes are designed such that in order to be successful, it is expected that students will spend a minimum of forty-five hours for each unit of credit (normally three hours per unit per week), including preparing for class, participating in course activities, completing assignments, and so on. More details about student workload can be found in University Policy S12-3 at http://www.sjsu.edu/senate/docs/S12-3.pdf.

Homework, Exams and Final Projects are expected for this class. Homework is due on Canvas by midnight on the due date. Each assigned problem requires a solution and an explanation (or work) detailing how you arrived at your solution. Cite any outside sources used to solve a problem. When grading an assignment, I may ask for additional information.

NOTE that University policy F69-24 at http://www.sjsu.edu/senate/docs/F69-24.pdf states that "Students should attend all meetings of their classes, not only because they are responsible for material discussed therein, but because active participation is frequently essential to insure maximum benefit for all members of the class. Attendance per se shall not be used as a criterion for grading."

✔ Grading Information

Semester grade will be computed as a weighted average of the scores obtained in each of the four categories listed above. No make-up tests or quizzes will be given, and no late homework (or other work) will be accepted. Also, in-class work must be completed in the section that you are enrolled in.

Criteria

Determination of Grades

Nominal Grading Scale:

Percentage	Grade
97 – 100 plus	A+
93 – 96	A
90 – 92	A-

87 – 89	B+
83 – 86	B
80 – 82	B-
77– 79	C+
73 – 76	C
70 – 72	C-
67 – 69	D+
63 – 66	D
60 - 62	D-
0-59	F

Breakdown

Homework: 25%

Exam 1: 25%

Exam 2: 25%

Final Project: 25%

Note that "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 F13-1 at http://www.sjsu.edu/senate/docs/F13-1.pdf 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

CS171 / Introduction to Machine Learning, Fall 2024, Course Schedule

The schedule is subject to change with fair notice communicated via Canvas course page

Course Schedule

Week	Date	Topics, Readings, Assignments, Deadlines
1	08/22	Introduction
1	08/27	Supervised Learning: Classification, Regression, Generalization and Model Complexity
2	08/29	Supervised Learning: K Nearest Neighbors, Linear Models, Naïve Bayes
2	08/30	Supervised Learning: Decision Trees, Ensembles DTs, Kernelized SVM
3	09/03	Supervised Learning: Neural Networks, Backpropagation and Optimization
3	09/05	Supervised Learning: Deep Learning with Dense Feedforward Neural Networks
4	09/10	Supervised Learning: Deep Learning with Convolutional Neural Networks
4	09/12	Supervised Learning: Deep Learning with Recurrent Neural Networks
5	09/17	Supervised Learning: Discussions on Convolution and Recurrence-driven Models

Week	Date	Topics, Readings, Assignments, Deadlines
5	09/19	Unsupervised Learning and Preprocessing: Principal Components Analysis, Non-negative Matrix Factorization
6	09/24	Unsupervised Learning and Preprocessing: Manifold Learning with t-SNE, UMAP
6	09/26	Unsupervised Learning and Preprocessing: K-Means, K-Means++, Agglomerative Clustering, DBSCAN
7	10/01	Representing Data and Engineering Features: Categorical Variables, Binning, Discretization, Linear Models and Trees, Interactions, Polynomials, Automatic Feature Selection, Cross Validation
7	10/03	Best practices in Machine Learning Operations (MLOPs)
8	10/08	Computer Vision: Object Detection, Boundary Detection, Image Segmentation
8	10/10	Mathematical Optimization: Deterministic Search and Guided Random Search
9	10/15	Working with Structured Data, Machine Learning in the Frequency Domain
9	10/17	Introduction to Recommender Systems
10	10/22	Exam 1
10	10/24	Generative Modeling: Variational Autoencoders, Autoregressive Models
11	10/29	Generative Modeling: Generative Adversarial Networks, Diffusion Models
11	11/31	Working with Text Data: Types of Data, Representation as Bag-of-Words

Week	Date	Topics, Readings, Assignments, Deadlines
12	11/05	Working with Text Data: Stopwords, Rescaling, Model Coefficients, n-Grams, Advanced Tokenization, Stemming and Lemmatization
12	11/07	Generative AI : BERT, Transformers, Generative Pre-trained Transformers
13	11/12	Large Language Models: Introduction
13	11/14	Large Language Models: Training, Inference, Prompt Engineering, Interpretability, Reinforcement Learning through Human Feedback (RLHF)
14	11/19	Reinforcement Learning: Introduction, Tabular Solution Methods
14	11/21	Reinforcement Learning: Approximate Solution Methods, Course Wrap Up
15	11/26	Exam 2
15	11/28	No Class
16	12/03	Industry Speaker Session
16	12/05	Project Presentations
16	12/10	Project Presentations