San José State University
Department of Computer Science
CS171, Introduction to Machine Learning, Section 3, Fall, 2022

Course and Contact Information

Instructor:	Saptarshi Sengupta, PhD
Office Location:	Duncan Hall 239
Telephone:	408-924-4808
Email:	saptarshi.sengupta@sjsu.edu
Office Hours:	Thursday, 10:30 AM – 12:30 PM
Class Days/Time:	MW 10:30 AM - 11:45 AM
Classroom:	MacQuarrie Hall 225
Prerequisites:	CS 146 (with a grade of "C-" or better)

Course Format

Faculty Web Page 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 Description

The course is an in-depth exploration of the underlying mathematics as well as some practical applications in the field of machine learning. The course also focuses on designing algorithms that allow computers to learn and take actions in a probabilistic fashion using statistical inferences. Students will explore Probabilistic Reasoning, Belief Networks, Bayesian learning, Hidden Variable Learning, Supervised and Unsupervised Learning, Linear Modeling, Gaussian Processes, Mixture Models, Discrete State Markov Models. In addition, students will explore the fundamentals and various use cases in Deep Learning, Natural Language Processing and Reinforcement Learning using packages such as Scikit-Learn, Tensorflow, Keras, PyTorch, SpaCy, NLTK etc.

Course Learning Outcomes (CLO)

The focus of this course will be machine learning, with examples from fields such as computer vision, anomaly detection, health informatics and prognostics. 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 in the industry.

Required Texts/Readings

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

Other technology requirements / equipment / material

Python 3, Scikitlearn libraries, numpy/scipy, tensorflow/keras, gym, Jupyter notebooks. Installing Anaconda is highly recommended. I will be using Jupyter Notebook in my code demos in class.

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."

CoS COVID-19 and Monkeypox Safety

Students registered for a College of Science (CoS) class with an in-person component should view the CoS COVID-19 and Monkeypox Training slides for updated CoS, SJSU, county, state and federal information and guidelines, and more information can be found on the SJSU Health Advisories website. By working together to follow these safety practices, we can keep our college safer. Failure to follow safety practice(s) outlined in the training, the SJSU Health Advisories website, or instructions from instructors, TAs or CoS Safety Staff may result in dismissal from CoS buildings, facilities or field sites. Updates will be implemented as changes occur (and posted to the same links).

Final Examination or Evaluation

The final project presentations will be administered in person.

Grading Information

- 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.

Determination of Grades

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.

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

Classroom Protocol

- **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

University Policies (Required)

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Programs' Syllabus Information web page at http://www.sjsu.edu/gup/syllabusinfo/". Make sure to review these policies and resources.

CS171 / Introduction to Machine Learning, Fall 2022, 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
2	08/24	Why Machine Learning
2	08/29	Supervised Learning: Classification, Regression, Generalization and Model Complexity
3	09/31	Supervised Learning: K Nearest Neighbors, Linear Models, Naïve Bayes
3	09/05	**Labor Day (Campus Closed)**
4	09/07	Supervised Learning: Decision Trees, Ensembles DTs, Kernelized SVM
4	09/12	Supervised Learning: Neural Networks Fundamentals
5	09/14	Supervised Learning: Deep Learning
5	09/19	Supervised Learning: Deep Learning with Convolutional Neural Networks
6	09/21	Supervised Learning: Deep Learning with Recurrent Neural Networks
6	09/26	Supervised Learning: Some Applications of Deep Learning
7	09/28	Unsupervised Learning and Preprocessing: Principal Components Analysis
7	10/03	Unsupervised Learning and Preprocessing: Non-negative Matrix Factorization
8	10/05	Unsupervised Learning and Preprocessing: Manifold Learning with t-SNE
8	10/10	Unsupervised Learning and Preprocessing: K-Means and K-Means++
9	10/12	Unsupervised Learning and Preprocessing: Agglomerative Clustering, DBSCAN
9	10/17	**Exam 1**
10	10/19	Representing Data and Engineering Features: Categorical Variables
10	10/24	Representing Data and Engineering Features: Binning, Discretization, Linear Models and Trees
11	10/26	Representing Data and Engineering Features: Interactions, Polynomials, Automatic Feature Selection
11	10/31	Model Evaluation and Improvement: Cross Validation
12	11/02	Algorithmic Chains and Pipelines
12	11/07	Working with Text Data: Types of Data, Representation as Bag-of-Words

Week	Date	Topics, Readings, Assignments, Deadlines
13	11/09	Working with Text Data: Stopwords, Rescaling, Model Coefficients
13	11/14	Working with Text Data: n-Grams, Advanced Tokenization, Stemming and Lemmatization
14	11/16	**Exam 2**
14	11/21	Reinforcement Learning: Introduction
15	11/23	**Non-instructional Day (No Classes)**
15	11/28	Reinforcement Learning: Tabular Solution Methods
16	11/30	Reinforcement Learning: Approximate Solution Methods, Course Wrap Up
16	12/05	**Project Presentations:** Monday, December 05, 10:30 AM – 11:45 AM