

C S 12A: INTRODUCTION TO MACHINE LEARNING

Foothill College Course Outline of Record

Heading	Value
Effective Term:	Summer 2025
Units:	4.5
Hours:	4 lecture, 2 laboratory per week (72 total per quarter)
Prerequisite:	C S 3A.
Advisory:	Students will benefit from prior exposure to statistics.
Degree & Credit Status:	Degree-Applicable Credit Course
Foothill GE:	Non-GE
Transferable:	CSU/UC
Grade Type:	Letter Grade (Request for Pass/No Pass)
Repeatability:	Not Repeatable

Student Learning Outcomes

- Explain the difference between the supervised, unsupervised, and reinforcement learning paradigms, and define key algorithms and models within each paradigm
- Use Python packages to train machine learning models and evaluate their performance using metrics appropriate to the task (e.g., accuracy, precision, recall)
- Understand how dataset quality can contribute to or detract from equitable model outcomes, and the ramifications of these outcomes on society

Description

A survey of machine learning algorithms and modern packages. Includes models in supervised, unsupervised, and reinforcement learning. Explores the entire machine learning pipeline from dataset selection through model evaluation. Students will gain practical experience coding models, with less emphasis on math and theory.

Course Objectives

The student will be able to:

1. Explain the role of statistics and linear algebra in artificial intelligence
2. Use Python packages such as Pandas and Matplotlib to explore datasets prior to implementing a machine learning model
3. Describe the difference between regression and classification
4. Implement data set splitting and cross-validation and explain the value of these practices
5. Evaluate model performance, and explain underfitting, overfitting, and the bias-variance tradeoff
6. Independently implement at least one model in each of supervised, unsupervised, and reinforcement learning
7. Use Python packages such as scikit-learn, Keras, TensorFlow, and PyTorch to develop machine learning models

8. Discuss security, ethics, and equity in the context of machine learning, including explainability, accountability, and interpretability

Course Content

1. Mathematical foundations
 - a. Statistics
 - b. Vectors
 - c. Matrices
2. Preparing data
 - a. Explore datasets and formulate a question that machine learning can solve
 - b. Visualizing
 - c. Cleaning
 - d. Encoding
 - e. Curse of dimensionality
 - f. PCA
 - g. Data pitfalls
 - h. Noise
3. Techniques of regression and classification
 - a. Categorize machine learning problems as regression or classification
4. Preparing a dataset for validation
 - a. Split a dataset into a train set and a test set
 - b. Partition a dataset for k-fold cross validation
5. Error analysis and performance evaluation
 - a. RMSE
 - b. Bias and variance
 - c. Overfitting and underfitting
 - d. Confusion matrices
 - e. Precision, recall, and F1
6. Machine learning models
 - a. Supervised learning
 - i. Common features and techniques
 1. Hyperparameters
 2. Loss functions
 3. Gradient descent
 - ii. Linear regression
 - iii. Logistic regression
 - iv. K-nearest neighbor
 - v. Support vector machines
 - vi. Decision trees
 1. Entropy
 2. Information gain
 - vii. Ensemble learning
 1. Bagging
 2. Boosting
 - viii. Naïve Bayes
 - ix. Artificial neural networks
 - b. Unsupervised learning
 - i. Clustering
 - c. Reinforcement learning
 - i. Markov decision process
 - ii. Q-learning
 - d. Deep learning

7. Machine learning packages
 - a. Installation and importing
 - b. Documentation and API
 - c. Machine learning pipeline
8. Safety and ethics of machine learning
 - a. Accountability
 - b. Interpretability
 - c. Explainability
 - d. Disproportionate impact
 - e. Historical examples of racism and other harmful bias in machine learning
 - f. Preventing racism and other harmful bias in machine learning

Lab Content

1. Familiarity with package installation
 - a. Installation
 - b. Importing
 - c. Documentation
2. Exploration of data table and graphing packages
 - a. Identify and address incomplete data
 - b. Find key statistics from data tables
 - c. Generate a pair-plot to visually explore co-variance
3. Prepare a dataset for machine learning
 - a. Identify and address missing information
 - b. Identify extraneous columns
 - c. Quantify qualitative data where possible and necessary
 - d. Perform binary or class encoding
 - e. Apply function transformations to normalize data
 - f. Apply principal component analysis to reduce dimensionality
 - g. Split data into testing and training sets
 - h. Split data into k-fold sets
4. Use machine learning packages to develop and test at least one model in each domain:
 - a. Supervised learning
 - b. Unsupervised learning
 - c. Reinforcement learning
5. Use Python packages or student code to produce error reports
 - a. RMSE
 - b. Confusion matrix
 - c. Precision, recall, and F1

Special Facilities and/or Equipment

1. The college will provide access to a computer laboratory with Python and an IDE installed, with sufficient privileges to allow students to install Python packages.
2. The college will provide a website or course management system with an assignment posting component (through which all lab assignments are to be submitted) and a forum component (where students can discuss course material and receive help from the instructor). This applies to all sections, including on-campus (i.e., face-to-face) offerings.
3. When taught online, the college will provide a fully functional and maintained course management system through which the instructor and students can interact.
4. When taught online, students must have currently existing email accounts and ongoing access to computers with internet capabilities.

Method(s) of Evaluation

Methods of Evaluation may include but are not limited to the following:

Tests and quizzes

Lab notebook

Written laboratory assignments which include source code, sample runs, and documentation

Reflective papers

Final examination or project

Method(s) of Instruction

Methods of Instruction may include but are not limited to the following:

Instructor-authored lectures which include mathematical foundations, theoretical motivation, and coding implementation of machine learning algorithms

Detailed review of assignments which includes model solutions and specific comments on the student submissions

Discussion which engages students and instructor in an ongoing dialog about machine learning

Instructor-authored labs that rigorously demonstrate a student's ability to implement machine learning models

Representative Text(s) and Other Materials

Géron, Aurélien. [Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow](#). 2022.

Raschka, Sebastian, Yuxi (Hayden) Liu, and Vahid Mirjalili. [Machine Learning with PyTorch and Scikit-Learn](#). 2022.

Types and/or Examples of Required Reading, Writing, and Outside of Class Assignments

1. Reading
 - a. Textbook assigned reading averaging 30 pages per week
 - b. Reading the supplied handouts and modules averaging 10 pages per week
 - c. Reading online resources as directed by instructor though links pertinent to programming
 - d. Reading library and reference material directed by instructor through course handouts
2. Writing
 - a. Writing technical prose documentation that supports and describes the programs that are submitted for grades

Discipline(s)

Computer Science