Oregon State University
Ecampus

Course Name: Applied Machine Learning
Course Number: CS 519-400
Credits: 4
Instructor name: Professor Liang Huang
Instructor email: liang.huang@oregonstate.edu

Teaching Assistants name and contact info:

Sizhen Li lisiz@oregonstate.edu
Junkun Chen chenjun2@oregonstate.edu

Course Description

Explores Machine Learning Basics (variance and bias, underfitting and overfitting, etc). Reviews Linear Algebra and Numpy. Examines k-Nearest Neighbors, Linear Classification (Perceptron and Online Learning), and Linear and Non-Linear Regression. Explores applications in Housing Price Prediction (Kaggle Contest) and Text Classification (Sentiment Analysis).

PREREQS: Graduate standing.

Communication

Please post all course-related questions on the forum specified by the instructor so that the whole class may benefit from our conversation. Please contact us privately only for matters of a personal nature; we will not reply to any technical questions via email. We will strive to return your assignments and grades for course activities to you within seven days of the due date.

Course Credits

This course combines approximately 120 hours of instruction, online activities, and assignments for 4 credits. Each week a student is expected to spend ~2 hours watching lecture videos, ~2 hours reading the textbook, ~2 hour studying the slides or other materials released by the instructor, and ~6 hours on homework.

Technical Assistance

If you experience any errors or problems while in your online course, contact 24-7 Canvas Support through the Help link within Canvas. If you experience computer difficulties, need help downloading a browser or plug-in, or need assistance logging into a course, contact the IS Service Desk for assistance. You can call (541) 737-8787 or visit the IS Service Desk online.

Learning Resources

Textbook: A Course in Machine Learning (CIML, freely available online), by Hal Daume III.
http://ciml.info/

Canvas

Oregon State University
Ecampus

This course is delivered via Canvas, where you will interact with your classmates and your instructor. You will access the learning materials within the course site, such as the syllabus, class discussions, assignments, projects, and quizzes. To preview how an online course works, visit the Ecampus Course Demo. For technical assistance, please visit Ecampus Technical Help.

Measurable Student Learning Outcomes

1. Students will be able to formulate the components of a machine learning algorithm.
2. Students will be able to contrast training, test, and generalization errors, to identify and interpret underfitting and overfitting, and to use methods to cope with underfitting and overfitting.
3. Students will be able to formulate and implement a k-NN classifier.
4. Students will be able to formulate and implement the averaged perceptron classifier.
5. Students will be able to interpret and extend the perceptron convergence proof.
6. Students will be able to use linear regression in a real-world prediction task.
7. Students will be able to use support vector machines and kernels in a real-world classification task.
8. Students will be able to use linear classifiers in real-world text classification and sentiment analysis tasks.

Evaluation of Student Performance

Background survey (on Canvas): each student gets 2% by submitting.
Quizzes 1 & 2 (on Canvas), 10% + 8% = 18%.
Programming Assignments (HW1-HW4), 20% + 15% + 15% + 15% = 65%.
Paper Review (HW5), 15%.

Course Content

Week	Topics	Readings	Learning Activities
Unit 1: Introduction, k-NN, and Linear Algebra/Numpy Review			
1	Introduction Underfitting/Overfitting Training, Testing, & Generalization Errors *k*-Nearest Neighbors Classification Decision trees	CIML 0 CIML 2.4 CIML 2.5 CIML 3 CIML 1	Lecture Videos Slides Background survey Quiz 1 (ML basics)
2-3	Geometric Review of Linear Algebra Numpy Tutorial (also matplotlib) Data preprocessing in machine learning	Extra readings Jupyter notebook	Lecture Videos Slides Quiz 2 (linear algebra) HW1: k-NN: income>50k?
Unit 2: Linear Classification and Perceptron Algorithm			
4	Linear Classifiers Perceptrons (1)	CIML 4 CIML 5	Lecture Videos Slides

Week	Topics	Readings	Learning Activities
5	Perceptrons (2) Perceptron vs. Logistic Regression	CIML 4 CIML 5 CIML 9.6	Lecture Videos Slides HW2 (same data as HW1)
Unit 3: Linear and Polynomial Regression			
6-7	Linear and Polynomial Regression	CIML 7.6	Lecture Videos Slides HW3: housing price prediction (Kaggle competition)
Unit 4: Applications: Text Classificaiton			
8-9	Application: Text Classification; Sentiment Analysis	Not in CIML	HW4: sentiment analysis
Unit 5: Exposure to Cutting-Edge ML Research			
10	Paper Review	Suggested papers	HW5: paper review

Course Policies

Grading Policies

Quizzes (on Canvas): you have two (2) chances for each quiz; graded automatically. Testing basic concepts of machine learning and linear algebra.

Programming Assignments (HW1-HW4): graded by correctness and relative ranking of prediction accuracy, and by clarity of the report. Implementations should be done in Python+numpy only. No IDEs are needed. We assume a Unix-like environment such as Linux or Mac OS X; note that Windows are not supported – if you do not have a Mac or Linux machine, you can use OSU School of Engineering servers: `ssh access.engr.oregonstate.edu`; if you do not have an ENGR account, please register here: https://it.engineering.oregonstate.edu/get-engr-account.

You're highly recommended to set up SSH keys so that you don't need to type your password or go through Duo authentication when connecting to the ENGR servers: https://it.engineering.oregonstate.edu/ssh-keygen

Paper Review (HW5): graded against the paper review template.

Late Work Policy

Each student can be late by 24 hours only once without penalty. No more late submissions will be accepted.

Incompletes

Incomplete (I) grades will be granted only in emergency cases (usually only for a death in the family, major illness or injury, or birth of your child), and if the student has turned in 80% of the points possible (in other words, usually everything but the final paper). If you are having any

difficulty that might prevent you completing the coursework, please don't wait until the end of the term; let me know right away.

Guidelines for a Productive and Effective Online Classroom

Students are expected to conduct themselves in the course (e.g., on discussion boards, email) in compliance with the university's regulations regarding civility. Civility is an essential ingredient for academic discourse. All communications for this course should be conducted constructively, civilly, and respectfully. Differences in beliefs, opinions, and approaches are to be expected. In all you say and do for this course, be professional. Please bring any communications you believe to be in violation of this class policy to the attention of your instructor.

Active interaction with peers and your instructor is essential to success in this online course, paying particular attention to the following:

- Unless indicated otherwise, please complete the readings and view other instructional materials for each week before participating in the discussion board.
- Read your posts carefully before submitting them.
- Be respectful of others and their opinions, valuing diversity in backgrounds, abilities, and experiences.
- Challenging the ideas held by others is an integral aspect of critical thinking and the academic process. Please word your responses carefully, and recognize that others are expected to challenge your ideas. A positive atmosphere of healthy debate is encouraged.

Statement Regarding Students with Disabilities

Accommodations for students with disabilities are determined and approved by Disability Access Services (DAS). If you, as a student, believe you are eligible for accommodations but have not obtained approval, please contact DAS immediately at 541-737-4098 or at http://ds.oregonstate.edu. DAS notifies students and faculty members of approved academic accommodations and coordinates implementation of those accommodations. While not required, students and faculty members are encouraged to discuss details of the implementation of individual accommodations.

Accessibility of Course Materials

All materials used in this course are accessible. If you require accommodations please contact Disability Access Services (DAS).

Additionally, Canvas, the learning management system through which this course is offered, provides a vendor statement certifying how the platform is accessible to students with disabilities.

Expectations for Student Conduct

Student conduct is governed by the university's policies, as explained in the Student Conduct Code. Students are expected to conduct themselves in the course (e.g., on discussion boards, email postings) in compliance with the university's regulations regarding civility.

Academic Integrity

Students are expected to comply with all regulations pertaining to academic honesty. For further information, visit Student Conduct and Community Standards, or contact the office of Student Conduct and Mediation at 541-737-3656.

OAR 576-015-0020 (2) Academic or Scholarly Dishonesty:

a) Academic or Scholarly Dishonesty is defined as an act of deception in which a Student seeks to claim credit for the work or effort of another person, or uses unauthorized materials or fabricated information in any academic work or research, either through the Student's own efforts or the efforts of another.
b) It includes:
    i) CHEATING - use or attempted use of unauthorized materials, information or study aids, or an act of deceit by which a Student attempts to misrepresent mastery of academic effort or information. This includes but is not limited to unauthorized copying or collaboration on a test or assignment, using prohibited materials and texts, any misuse of an electronic device, or using any deceptive means to gain academic credit.
    ii) FABRICATION - falsification or invention of any information including but not limited to falsifying research, inventing or exaggerating data, or listing incorrect or fictitious references.
    iii) ASSISTING - helping another commit an act of academic dishonesty. This includes but is not limited to paying or bribing someone to acquire a test or assignment, changing someone's grades or academic records, taking a test/doing an assignment for someone else by any means, including misuse of an electronic device. It is a violation of Oregon state law to create and offer to sell part or all of an educational assignment to another person (ORS 165.114).
    iv) TAMPERING - altering or interfering with evaluation instruments or documents.
    v) PLAGIARISM - representing the words or ideas of another person or presenting someone else's words, ideas, artistry or data as one's own, or using one's own previously submitted work. Plagiarism includes but is not limited to copying another person's work (including unpublished material) without appropriate referencing, presenting someone else's opinions and theories as one's own, or working jointly on a project and then submitting it as one's own.
c) Academic Dishonesty cases are handled initially by the academic units, following the process outlined in the University's Academic Dishonesty Report Form, and will also be referred to SCCS for action under these rules.

Student Evaluation of Courses

The online Student Evaluation of Teaching system opens to students the Wednesday of week 8 and closes the Sunday before Finals Week. Students will receive notification, instructions and the link through their ONID. They may also log into the system via Online Services. Course evaluation results are extremely important and used to help improve courses and the learning experience of future students. Responses are anonymous (unless a student chooses to "sign" their comments agreeing to relinquish anonymity) and unavailable to instructors until after grades have been posted. The results of scaled questions and signed comments go to both the instructor and their unit head/supervisor. Anonymous (unsigned) comments go to the instructor only.