Machine Learning, Fall 2022

Time: MWF 12:30pm-1:20pm

Location: Felgar Hall 0304

Instructor: Chao Lan (clan@ou.edu)

Teaching Assistant: Yan He (heyan@ou.edu)

Office Hours:

– W 1:30pm-4:30pm, Dr. Lan, DEH 341 (later move to 210-E)

– F 3pm-4:30pm, Yan, Virtual Office Hour, Zoom ID: 8704765409, PWD: machineL

---

1. Course Description

This course introduces the fundamentals of machine learning. It will cover

(1) Regression models, their learning techniques and probabilistic interpretations. Subjects include least square, ridge, Lasso, MLE, MAP, kernel regression.

(2) Dimensionality reduction techniques. Subjects include general feature selection, principal component analysis, canonical correlation analysis.

(3) Classification models and their learning techniques. Subjects include logistic regression, naive Bayes, k nearest neighbor, decision tree, support vector machine, neural network.

(4) Clustering. Subjects include K-means.

(5) Ensemble methods. Subjects include bagging (random forest) and boosting (adaboost).

(6) Other topics if time permits.

Students are expected to understand both intuitions and mathematics of the lectured subjects and be able to derive some from scratch, which will be evaluated based on written assignments and exams. Students are also expected to implement most lectured techniques from scratch in Python (i.e., implement logistic regression by implementing the gradient decent optimization technique instead of calling the logistic regression function in Python libraries), which will be evaluated based on programming assignments.

2. Reference

T. Hastie, R. Tibshirani and J. Friedman. The Elements of Statistical Learning. Springer, 2009.

C. Bishop. Pattern Recognition and Machine Learning. Springer, 2006.

S. Boyd and L.Vandenberghe. Convex Optimization. Cambridge University Press, 2004.

L. Wasserman. All of Statistics. Springer, 2004.

** All reference books are available online.*

3. Assignments and Exams

There will be three exams, 5 to 10 assignments and some quizzes. The final points will have 45% from the exams, 50% from the assignments and 5% from the quizzes. There will be no bonus points.

If an assignment is submitted within one week after the due, only 80% of its earned points will be counted towards the final points. If it is submitted one week after the deadline, it will not be counted.

Students cannot collaborate on any exam. Students can collaborate on the assignments but need to clarify the collaborator in submission and be ready to independently defend their own answers upon request.

4. Expected Background Knowledge

Basic knowledge on probability, statistics, linear algebra and vector calculus are expected to fully understand most lectured subjects and complete written assignments and exams.

Sufficient programming skills (in Python) are expected to complete programming assignments on time.

5. Other Requirements

Attendance is not required. Outline of the lecture notes will be posted on Canvas.

Students are encouraged to continue wearing masks in class and during office hours.