EE2211 Introduction to Machine Learning

Lec 0

Wang Xinchao and Vincent Tan

Welcome to EE2211!

- Team
  - Lecturers
    - Xinchao Wang (Lec 1-3, Lec 10-12)
    - Vincent Tan (Lec 4-9)

  - Python Tutor
    - Xingyi Yang (Week 1-2)

  - Graduate Assistants (i.e., Graders)
    - Songhua Liu
    - Xingyi Yang
    - Weihao Yu
    - Gongfan Fang
    - Xinyin Ma

  - Support and Coordinators
    - Celine Cheong
    - Jingwen Ye

  - Tutors
    - Prof. Tham Chen Kong (T12)
    - Christopher Moy Shin Lee Lan Chong (T03, T07)
    - Goh Shu Ting (T01, T05)
    - Henry Tan (T19, T21, T25)
    - Qingqing Ni (T11, T16, T27)
    - Abhijit Singh (T10, T18, T26)
    - Erik Maurits Spaans (T09, T24)
    - Koo Wei De (T05, T28)
    - Matsutake Teppei (T06, T08, T17)
    - Pan Jiachun (T02, T23)
    - Ricky Wahyudi (T13, T14)
    - Tan Yue Feng (T04, T20)
    - Wu Yilei (T22)

Logistics

- Schedule
  - 12 Weeks Lectures, starting from Week 1
  - 12 Weeks Tutorials, starting from Week 2
  - 2 Programming Tutorials (optional and highly recommended)
    - Week 1 – 2, Friday (i.e., 13 Jan and 20 Jan), right after the lecture
  - 1 mid-term Quiz (using ExamSoft)
    - Held on <u>4 March 2023</u>, content up to Week 6 (inclusive)
  - 1 briefing session on ExamSoft
    - Held on <u>26 Jan 2023 (Week 3), 4 to 4:30PM</u>
  - 1 Final Exam (using ExamSoft)
    - Date to be confirmed by CDE
  - 3 Assignments
    - Assignment 1: released on Week 4, due on Week 6 (tentatively)
    - Assignment 2: released on Week 6, due on Week 9 (tentatively)
    - Assignment 3: released on Week 9, due on Week 13 (tentatively)

Logistics

- 3 Assignments (36%) + Tutorial Attendance (4%)
- 1 Mid-term (30%)
- 1 Final Exam (30%)

- Held online:
  – Lectures

- Held offline (in classrooms):
  – Tutorials

- Videos of lectures are made available after lectures.

Responsibility of Team Members

- All members, together, strive to serve you well! However, we have a huge class of >600 students!

- The lecturers will spare no effort in helping you, but it wouldn't be possible for us two to answer all questions from 600 students on time…

- Therefore, to get the most prompt and high-quality answers to your questions, when you have:
  - Logistic-related Questions, go to <u>Lecturers</u>
  - Lecture-related Questions, go to <u>Lecturers</u>
  - Fundamental Python Questions (Week 1-2), go to <u>GAs</u>
  - Tutorial-related Questions, go to <u>Tutors</u>
  - Assignment-related Questions, go to <u>GAs</u>

- We will also actively use **Canvas Discussion** to answer questions so that everyone benefits! Feel free to post questions there!

Reference Books

- [Book1] Andriy Burkov, "The Hundred-Page Machine Learning Book", 2019. (read first, buy later: http://themlbook.com/wiki/doku.php)

- [Book2] Andreas C. Muller and Sarah Guido, "Introduction to Machine Learning with Python: A Guide for Data Scientists", O'Reilly Media, Inc., 2017.

- [Book3] Jeff Leek, "The Elements of Data Analytic Style: A guide for people who want to analyze data", Lean Publishing, 2015.

- [Book4] Vincent Tan, "Introduction to Machine Learning for EE2211", https://vyftan.github.io/papers/ee2211book.pdf
  - Follows the flow of the lectures and contains many additional "theory" practice problems (no solutions yet)

Something to Note…

- The topic of machine learning, per se, is a mixture of concepts and applications. **To fully understand concepts, you have to code!**

- Hence,
  - During lecture, we focus on <u>teaching concepts</u>
    - Unfortunately, we won't be able to spend much time showing code since we will only have 2 hours (especially for Lecs 1-3 and Lecs 10-12)
  - During tutorials, we focus on <u>reviewing concepts and coding</u>
    - The tutors will discuss coding with you

- We understand that our students come from different departments all across CDE
  - Don't worry too much if you consider your coding skills to be not perfect, you will have chances to learn and improve in EE2211. ;-)
  - In past semesters, very majority of students end up doing great!

Course Contents

- Introduction and Preliminaries (Xinchao)
  - Introduction
  - Data Engineering
  - Introduction to Probability and Statistics
- Fundamental Machine Learning Algorithms I (Vincent)
  - Systems of linear equations
  - Least squares, Linear regression
  - Ridge regression, Polynomial regression
- Fundamental Machine Learning Algorithms II (Vincent)
  - Over-fitting, bias/variance trade-off
  - Optimization, Gradient descent
  - Decision Trees, Random Forest
- Performance and More Algorithms (Xinchao)
  - Performance Issues
  - K-means Clustering
  - Neural Networks