

# Machine Learning

(機器學習)

Course Introduction, 09/14/2015

Hsuan-Tien Lin (林軒田)

htlin@csie.ntu.edu.tw

Department of Computer Science  
& Information Engineering

National Taiwan University  
(國立台灣大學資訊工程系)



# Four Reasons for NOT Taking the Course (1/4)

## Complicated Contents

- from a Taiwanese student taking MIT ML class (translated):  
*The professor started writing math equations as if he was using some writing accelerator. After class I always felt feeble. The worst part is: I needed to understand the contents as soon as I can. Otherwise I cannot finish the homework and cannot follow up in the next class.*
- NTU ML class: designed to be **as good as** the best classes in the world
- similar things will happen to you

If you are not willing to be so **miserable**, ...

## Four Reasons for NOT Taking the Course (2/4)

### Strict Instructor

- Will you give me a second chance if I copy homework from other people? **NO**.
- Could you let me pass because I will be kicked out by the 1/2 rule? **NO**.
- Will you change my score from *F* to *C*? **NO**.
- How many will pass? **Any**, if necessary.

If you do not like a **strict** instructor, ...

## Four Reasons for NOT Taking the Course (3/4)

### Huge Loads

- from a student taking ML 2010 (posted on BBS):

lxxxxxx9: 作業光一小題就要我們test 100次?(  $100 * 10\text{min} = 16\text{hr}$ ) 唉 反覆檢查許多遍 希望是我的code寫壞了 不然出這作業的人真的很沒良心==

- our class: **four to six** times harder than a normal one in NTU
- around **eight** homework sets (and a hard final project)
- homework due within two weeks
- even have homework 0 **NOW**
  - already hard
  - though no need to submit

If you do not want to spend **so much time on homework**, ...

## Four Reasons for NOT Taking the Course (4/4) Online-learnable

- invited by NTU as two of the **Massive Online Open Courses** on NTU-Coursera: **Machine Learning Foundations** and **Machine Learning Techniques**
  - **slides** teaching
  - **Mandarin** teaching
  - **MOOC-synced** teaching
    - <https://class.coursera.org/ntumlone-003>
  - homework setting: multiple-choice problems **plus** detailed arguments
  - **“recorded” teaching mode**
- much **easier** to just learn online at home
  - you can **choose to only take the online courses instead**

If you do not want be **restricted by physical class, ...**

from a student in ML2013 (final feedback):

人活的好好的，爲什麼一定要修Machine Learning呢？XD 這是一門體驗各種崩潰、絕望的課程，人生能被課程電成這樣可能也就這麼一回

May the Brave Ones Stay

## Basic Information

- instructor: Hsuan-Tien Lin ([htlin@csie.ntu.edu.tw](mailto:htlin@csie.ntu.edu.tw))
- office hour: after class or by appointment
- course webpage (CEIBA):  
<https://ceiba.ntu.edu.tw/1041mlearn>  
announcements, homework, reference handouts, etc.
- mailing list: supported by CEIBA
- discussion forum: supported by **ML-Foundations on Coursera**  
—<https://class.coursera.org/ntumlone-003>
- new: **four-hour credits** instead of three-hour: as suggested by previous students

**update your secondary email address on  
CEIBA and register for ML-Foundations**

# Enrollment

- **With school policy, students who have taken the three-hour version is NOT allowed to take this class again!**
- at most 176 in Room 103 of CSIE Building
- target size 220, which is 125% of 176—assuming that 1/5 will not show up in class
- priority: CSIE (CSIE, INM) > EECS > others
- auditing: welcomed (to sit) only if there is an empty chair

**Leave as soon as possible! Give your classmates a chance to be miserable.**



# Teaching Assistants

Kuan-Hao Huang, You-Lin Tsou, Hong-Min Chu, Hsien-Chun Chiu,  
Liang-Wei Chen, Yu-An Chung, Meng-Yuan Yang, Yao-Yuan Yang,  
Si-An Chen

- forum for course/homework material questions: on Coursera
- email for grading questions **only**: [m12015ta@csie.ntu.edu.tw](mailto:m12015ta@csie.ntu.edu.tw)
- TA Hour for more interactive discussions: **to be announced**

**To save TA loads, questions about course/homework materials will only be answered on the forum and/or TA hours after the forum starts running.**

# THE Book

## *Learning from Data: A Short Course*

Y. Abu-Mostafa (Caltech), M. Magdon-Ismail (RPI), H.-T. Lin (NTU)

- idea initiated during 2008
- 5 chapters, closely needed for the first half of the class
- other e-Chapters to be used in the second half of the class
- teaching with the book and suggested reading within the book

## Getting the Book to Read

- NTU Library: one reserved copy in the shared course material area
- R536: some shared copies to be read in the room
- Chuan-Hwa Book Company: imported some limited copies of the book
  - Ms. Jen Huang ([jen@chwa.com.tw](mailto:jen@chwa.com.tw)) at 0958-008-962
  - may or may not offer group discounts
- Amazon: main selling channel in the US, but can be expensive/slow for international shipping
  - <http://www.amazon.com/gp/product/1600490069>
- Bulk order from U.S.: secondary selling channel, usually takes two weeks to arrive — <http://amlbook.com>
- **If the book is not affordable to you but you really want to read it: email me ([htlin@csie.ntu.edu.tw](mailto:htlin@csie.ntu.edu.tw)) and I'll see how I can help.**

# Getting Future Draft Chapters to Read

*mechanism: to be announced when needed*

## Your Privileges

- learn from the **first draft**
- obtain the draft **freely**

## Your Responsibilities

- discuss **actively** with me to improve the draft
- **do not distribute the draft**

**enrolling in this class means agreeing to  
the items above**

# THE Principle

Taking any unfair advantages over other class members is not allowed. It is everyone's responsibility to maximize the level of fairness.

- eating? fine, but no smells and no noise
- sleeping? fine, but no snoring
- cellphone? fine, but silent mode, and speak outside
- ...

**applies to instructor, TAs, students**

# Honesty

**NO CHEATING**

**NO LYING**

**NO PLAGIARISM**

**NO PIRATING of THE BOOK**

**very serious consequences**

# Grade

- no midterm, no final
- main reference: homework sets, final project
- raw score goes through some order-preserving normalization steps
  - raw score 80 with term rank A: possible
  - raw score 80 with term rank B: possible
  - raw score 60 with term rank F: possible
  - raw scores 80, 60 with term scores B, B: possible, but unlikely
  - raw scores 80, 60 with term scores F, B: **impossible**

# Collaboration and Open-Book

- homework discussions: encouraged
- but fairness?  
**write the final solutions alone and understand them fully**
- references (books, notes, Internet):  
consulted, but **not copied from**
- no need to lend/borrow solutions

to maximize fairness (everyone's responsibility),  
**lending/borrowing not allowed**



# Collaboration and Open-Book

to maximize fairness (everyone's responsibility),  
**lending/borrowing not allowed**

to maximize fairness (everyone's responsibility),  
**lending/borrowing not allowed**

to maximize fairness (everyone's responsibility),  
**lending/borrowing not allowed**

Deal? If your classmate wants to borrow homework from you, what do you say?

# Homework

- students: justify solutions clearly
- TAs: evaluate solutions fairly
- penalty for late parts:  
90% of value for 12-hour late, 80% one-day late, ...

**late homework should go to a box in R217**

# Programming Assignments

- about a third or half of the problems
- any programming language, any platforms
- upload **source code**, otherwise:  
10% of value only!
- **no sophisticated packages**

students' responsibility:

**ask TA in advance for what can/cannot be used**

# Important TODOs

- Update your secondary email address on CEIBA!
- Register ML-Foundations on Coursera!
- Do homework 0; send emails to TAs or post on Coursera forum for questions.

**May the Brave Ones Stay**

**Questions?**