

CSIE 5043: Machine Learning

Hsuan-Tien Lin

Dept. of CSIE, NTU

Course Introduction, 09/10/2012

nickname: Hsuan-Tien Machine Learning (HTML)

Only English

- English teaching
- English homework writing
- English email communications
- English forum discussions
- exception: Mandarin face-to-face discussions

If you are not comfortable with **English-teaching** classrooms, ...

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**, ...

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 (4/4)

Huge Loads

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

lxxxxxx9: 作業光一小題就要我們test 100次?(100*10min = 16hr) 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 seven homework sets (and a hard final project)
- homework due within two weeks
- even have homework 0 and 0.5 **NOW**
 - already hard
 - **no need to submit homework 0**, but need to do homework 0.5 immediately to follow up.

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

May the Brave Ones Stay

Basic Information

- instructor: Hsuan-Tien Lin (htlin@csie.ntu.edu.tw)
- office hour: Thursdays 14:20–15:20 or by appointment
- course webpage: <https://ceiba.ntu.edu.tw/1011ML>
announcements, homework, reference handouts, etc.
- mailing list: supported by CEIBA

update your secondary email address on CEIBA

- at most around 130 in Room 303 of Freshman Building
- readily have 98 now
- new cases: sign up for the univ. lottery first, may add **a few** more in the third week if space allows
- auditing: welcomed (to sit) only if there is an empty chair

Drop as soon as possible!

Give your motivated classmates a chance to be miserable.

Teaching Assistants

Ya-Hsuan Chang, Wei-Yuan Shen, Ching-Pei Lee,
Han-Jay Yang, Chun-Heng Huang

- forum for course/homework material questions: to be introduced later
- email for grading questions **only**: `m12012ta@csie.ntu.edu.tw`
- TA Hour for more interactive discussions:
Tuesdays 5:30pm to 6:30pm; Wednesdays 5:30pm to 6:30pm
— starting 9/18 and 9/19, with rooms to be announced soon

Go to TA hours to discuss with TAs and classmates!

Learning from Data: A Short Course

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

- idea initiated during HTML2008
- 5 chapters, closely needed for the first half of the class
- other draft chapters to be finalized, to be used in the second half of the class
- teaching with the book, many homework problems within the book, reading assignments within the book

Getting the Book to Read

- NTU Library: one reserved copy in the shared course material area
- Chuan-Hwa Book Company: imported some limited copies of the book
 - Ms. Jen Huang (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: email me (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**
- download 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

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

NO CHEATING

NO LYING

NO PLAGIARISM

NO PIRATING of THE BOOK

very serious consequences

- no midterm, no final
- main reference: homework sets, final project
- bonus 3%: participation in forum discussions

Participation in Forum Discussions

<http://book.caltech.edu/bookforum/index.php>

- forum of the book, with two boards about our class
- **1% bonus**: for introducing yourself on our board

<http://book.caltech.edu/bookforum/forumdisplay.php?f=144>

- **2% bonus**: if you have introduced yourself **and** your number of posts (everywhere in the forum) before the due date of the final project is within top 25% of class

To save TA loads, questions about course/homework materials will only be answered on the forum.

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?

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

four penalty-free half-days (gold medal) for deadline extension per person

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

Experimental Items for HTML2012

- two class slots
- bigger classroom
- published textbook for teaching (finally!)
- forum participation (bonus incentive and the forum-only policy)
- “watching” assignments (see homework 0.5)

Feel free to give us feedback (immediately or afterwards) to help improve the course!

Important TODOs

- Update your secondary email address on CEIBA!
- Sign the agreement form, register on the forum, and introduce yourself for the bonus.
- Do homework 0 and 0.5; post on the forum for questions.
- If you still want to be added, sign the form first (and try to enroll online) and wait for our decisions later.

May the Brave Ones Stay

Questions?