THE CHINESE UNIVERSITY OF HONG KONG
Print Course Catalog Details

Academic Org: Div of Computer Science & Engg – Subject: Computer Science

Course: CSCI5030 **Course ID:** 009698 **Eff Date:** 2024-07-01 **Crse Status: Active** **Apprv. Status:** Approved **[**New Course**]**
Machine Learning Theory 機器學習理論

This course first introduces fundamentals of machine learning with a large size of samples, including basic principles (maximum likelihood vs least redundancy) and typical structures (linear systems of hidden factors, mixture of local structures, and Markov temporal models), The second part of the course covers learning theories towards small sample size challenge, including major topics (model selection, learning regularization, two stage implementation, sparse learning, and automatic model selection) and three streams of efforts, namely generalization error estimation (CV, AIC, VC theory), shortest coding length (MML vs MDL) or similarly various Bayes (BIC, MAP, Laplace, marginal, and variational) , and BYY learning (BYY system, best harmony theory, Ying-Yang alternation updating, and five action circling implementation).

本科首先介紹大樣本下學習的基礎內容, 包括基本原理(最大似然擬合與最小冗餘之資訊保持)和典型結構(隱因數線性系統、局部結構之組合、馬爾可夫時序模型)。然後介紹有限樣本學習理論, 包括主要課題(模型選擇、規範學習、兩階段實現、稀疏學習、模型自動選擇)和三個方向上的理論探索,即估計泛化風險(CV, AIC, VC理論), 資料最短編碼(MML 與 MDL)及各種Bayes方法(BIC、MAP、Laplace、邊際、變分), 以及陰陽和諧學習(陰陽系統、和諧理論、與陰陽交替更新、五行循環執行)。

Grade Descriptor:

A

EXCELLENT – exceptionally good performance and far exceeding expectation in all or most of the course learning outcomes; demonstration of superior understanding of the subject matter, the ability to analyze problems and apply extensive knowledge, and skillful use of concepts and materials to derive proper solutions.

有關等級說明的資料，請參閱英文版本。

B

GOOD – good performance in all course learning outcomes and exceeding expectation in some of them; demonstration of good understanding of the subject matter and the ability to use proper concepts and materials to solve most of the problems encountered.

有關等級說明的資料，請參閱英文版本。

C

FAIR – adequate performance and meeting expectation in all course learning outcomes; demonstration of adequate understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

D

MARGINAL – performance barely meets the expectation in the essential course learning outcomes; demonstration of partial understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

F

FAILURE – performance does not meet the expectation in the essential course learning outcomes; demonstration of serious deficiencies and the need to retake the course.

有關等級說明的資料，請參閱英文版本。

Equivalent Offering:
Units: 3 (Min) / 3 (Max) / 3 (Acad Progress)
Grading Basis: Graded
Repeat for Credit: N
Multiple Enroll: N
Course Attributes: MSc Computer Science
MPhil-PhD Computer Sci & Erg

Topics:

COURSE OUTCOMES

Learning Outcomes:

Students will be able to understand fundamental concepts and develop critical thinking on :
1. ingredients, challenges, basic principles and a unified perspective of machine learning;
2. fundamentals of machine learning with a large size of samples;
3. efforts on learning theories and methods towards small sample size challenge;
4. a unified statistical learning framework: BYY learning and best harmony theory.

Course Syllabus:

This course first introduces fundamentals of machine learning with a large size of samples, including basic principles (maximum likelihood vs least redundancy) and typical structures (linear systems of hidden factors, mixture of local structures, and Markov temporal models), The second part of the course covers learning theories towards small sample size challenge, including major topics (model selection, learning regularization, two stage implementation, sparse learning, and automatic model selection) and three streams of efforts, namely generalization error estimation (CV, AIC, VC theory), shortest coding length (MML vs MDL) or similarly various Bayes (BIC, MAP, Laplace, marginal, and variational) , and BYY learning (BYY system, best harmony theory, Ying-Yang alternation updating, and five action circling implementation).

Assessment Type:

Others	: 25%
Short answer test or exam	: 65%
Selected response test or exam	: 10%

Feedback for Evaluation:

1. Mid-term course evaluation
2. Term-end course evaluation
3. Students' performance in their homework, the midterm exam and final exam

Required Readings:

To be provided by course teacher.

Recommended Readings:

Introduction to Machine Learning, Ethem ALPAYDIN, The MIT Press, 2004.
Machine Learning, Tom M. Mitchell, McGraw Hill, 1997.
Emerging themes on information theory and Bayesian approach, Xu, L, Li, Y D. eds,

special issue, Frontiers of Electrical and Electronic Engineering in China, 2010, 5(3).
Bayesian Ying-Yang system, best harmony learning, and five action circling. Xu, L.,
Journal of Frontiers of Electrical and Electronic Engineering in China, 5(3):281–328,2010.

OFFERINGS

1. CSCI5030 Acad Organization=CSEGV; Acad Career=RPG

COMPONENTS

LEC : Size=30; Final Exam=Y; Contact=3
TUT : Size=30; Final Exam=N; Contact=1

ENROLMENT REQUIREMENTS

1. CSCI5030 **Enrollment Requirement Group:**
For students in MSc Computer Science; or
For students in MPhil-PhD Computer Science & Engineering; or
For undergraduate students in Computer Science (CSCIU & CSCIN) or Computer Engineering (CENGU & CENGN)

Additional Information

VTL-Onsite face-to-face hrs 0
VTL-Online synch. hrs 0
VTL-Online asynch. hrs 0

< E N D O F R E P O R T >