MACHINE LEARNING

1.Thông tin về học phần (General Information)

Tên học phần (Course name): Machine Learning

Mã học phần (Course code): INT_E14121

Số tín chỉ (Number of credits): 3

Loại học phần (Course type): Compulsory

Học phần tiên quyết (Prerequisites):

- Algebra (BAS1201)
- Probability and Statistics (BAS1226)

Học phần trước (Previous courses):

Học phần song hành (Parallel courses):

Các yêu cầu đối với học phần (Course requirements):

- Lecture room: Projector, black board
- Laboratory:

Giờ tín chỉ đối với các hoạt động (Teaching and Learning hours):

- Lý thuyết (Lectures): 32h
- Bài tập (Exercises): 8h
- Bài tập lớn (Projects): 4h
- Thực hành (Labs): 0h
- Tự học (Individual reading): 1h

Địa chỉ Khoa/Bộ môn phụ trách học phần (Address of the Faculty/Department in charge of the course):

- Address: Faculty of Information Technology 1 - Posts and Telecommunications Institute of Technology, Km10, Nguyen Trai Street, Ha Dong District, Hanoi.
- Phone number: (024) 33510432

2. Mục tiêu học phần (Objectives)

Về kiến thức (Knowledge):

The goal of this course is to provide an introduction to machine learning algorithms and applications. The students will:

- Become familiar with fundamental concepts of machine learning and current achievements in this area.
- Gain knowledge of important machine learning algorithms including supervised learning (generative/discriminative learning, parametric/non-parametric learning, support vector machines); unsupervised learning (clustering, dimensionality reduction); and learning theory (bias/variance tradeoffs).
- Gain an application perspective of machine learning.

Kỹ năng (Skills):

Upon successful completion of this source, the students will be able to:

- Demonstrate the understanding of machine learning fundamental concepts, including potential applications.
- Understand and implement representative algorithms for supervised and unsupervised learning.
- Implement solutions for some real world applications using machine learning software or library.

Thái độ, Chuyên cần (Attitude):

Students are required to attend the classes and complete exercises and assignments.

3. Tóm tắt nội dung học phần (Description)

Machine learning explores the study and construction of algorithms that learn from data in order to make inferences about future outcomes. This course provides a broad introduction to algorithms and applications of machine learning. Topics include supervised learning (generative/discriminative learning, parametric/non-parametric learning, support vector machines); unsupervised learning (clustering, dimensionality reduction); and learning theory (bias/variance tradeoffs). The course will also discuss applications of machine learning, such as data mining, pattern recognition, and text and web data processing.

4. Nội dung chi tiết học phần (Outlines)

Chapter 1. Introduction
1.1. Machine learning definition and concepts
1.2. Types of learning
    1.2.1. Supervised learning
    1.2.2. Unsupervised learning
    1.2.3. Reinforcement learning
1.3. Components of a machine learning system
    1.3.1. Learning phase
    1.3.2. Prediction phase

Chapter 2. Linear regression
2.1. Simple linear regression
2.2. Multiple linear regression
2.3. Other considerations in the regression model
2.4. Project and Lab

Chapter 3. Logistic regression
3.1. An overview of classification
3.2. The logistic model
3.3. Estimating the regression coefficients
3.4. Making predictions
3.5. Multiple logistic regression
3.6. Logistic regression for multiple response classes
3.7. Project and lab

Chapter 4. Resampling Methods
4.1. Cross-validation
    4.1.1. The validation set approach
    4.1.2. Leave-One-Out Cross-validation
    4.1.3. k-Fold Cross-validation
    4.1.4. Bias-Variance Trade-off for k-Fold Cross-validation

4.1.5. Cross-Validation on classification problems
4.2. The bootstrap

Chapter 5. Tree-based methods
5.1. The basics of decision trees
5.1.1. Regression and classification trees
5.1.2. Tree versus linear models
5.1.3. Advantages and disadvantages of tree models
5.2. Bagging, random forests, boosting
5.2.1. Bagging
5.2.2. Random forests
5.2.3. Boosting
5.3. Project and Lab

Chapter 6. Support vector machines
6.1. Max-margin classifiers
6.2. Support vector classifiers
6.3. Support vector machines
6.4. SVMs with multiple classes
6.5. Project and Lab

Chapter 7. Principal component analysis
7.1. What are principal components
7.2. Interpretation of principal components
7.3. More on PCA
7.4. Other uses for principal components

Chapter 8. Clustering
8.1. K-Means clustering
8.2. Hierarchical clustering
8.3. Practical issues
8.4. Project and Lab

5. Học liệu (Textbooks)

5.1. Học liệu bắt buộc (Required Textbooks)

[1]. Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2014. *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated.

[2]. Shai Shalev-Shwartz and Shai Ben-David. 2014. *Understanding Machine Learning: From Theory to Algorithms.* Cambridge University Press, New York, NY, USA.

5.2. Học liệu tham khảo (Optional Textbooks)

[3]. Hal Daumé III. A Course in Machine Learning. http://ciml.info/.

[4]. A. Ng. *Lecture notes for Machine learning*. http://cs229.stanford.edu/materials.html

6. Phương pháp, hình thức kiểm tra – đánh giá kết quả học tập học phần (Grading Policy)

Grading method	Percentage	Group/Individual
- Attendance	10%	Individual
- Exercises	10%	Individual
- Mid-term projects/exam	10%	Group or individual

- Final examination	70%	Individual

Trưởng Bộ môn **(Head of Department)**	**Giảng viên biên soạn** **(Lecturer)**
Ngô Xuân Bách	**Từ Minh Phương**