City University of Hong Kong
Course Syllabus

offered by School of Data Science
with effect from Semester A 2019/20

Part I Course Overview

Course Title:	Statistical Machine Learning II
Course Code:	SDSC6001
Course Duration:	One Semester
Credit Units:	3
Level:	P6
Medium of Instruction:	English
Medium of Assessment:	English
Prerequisites: *(Course Code and Title)*	SDSC5001 Statistical Machine Learning I
Precursors: *(Course Code and Title)*	Nil
Equivalent Courses: *(Course Code and Title)*	Nil
Exclusive Courses: *(Course Code and Title)*	Nil

Part II Course Details

1. Abstract

This course focuses on the theoretical foundation and fundamental methods in unsupervised and supervised learning, including Support Vector Machines, Ensemble Methods, K-means, Spectral Clustering, Dimension Reduction, Regularization Methods, Neural Networks, and Deep learning methods as well as the discipline of applying Python to program and implement aforementioned algorithms and methods.

2. Course Intended Learning Outcomes (CILOs)

(CILOs state what the student is expected to be able to do at the end of the course according to a given standard of performance.)

No.	CILOs	Weighting (if applicable)	Discovery-enriched curriculum related learning outcomes (please tick where appropriate)		
			A1	*A2*	*A3*
1.	Understand principles and theories of machine learning and deep learning algorithms	20%	✓		
2.	Familiarize machine learning and deep learning methods and their computational algorithms	20%	✓		
3.	Understand the theories and criteria for model assessment, comparison, and selection	20%	✓	✓	
4.	Interpret modelling results via using machine learning and deep learning algorithms	20%	✓	✓	
5.	Implement taught machine learning methods to develop data-driven models via Python	20%		✓	✓
		100%			

A1: Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to self-life problems.

A3: *Accomplishments*
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

3. Teaching and Learning Activities (TLAs)

(TLAs designed to facilitate students' achievement of the CILOs.)

TLA	Brief Description	CILO No.						Hours/week (if applicable)
		1	2	3	4	5		
Lecture	Introduce principles of machine learning and deep learning algorithms covered in this course	✓	✓	✓	✓			26 hours/sem
Laboratory work	Assist students to develop the ability of implementing machine learning and deep learning algorithms to analyse data and develop data-driven models through lab activities		✓	✓	✓	✓		13 hours/sem

Lectures cover not only the narrowly focused techniques in engineering economy but also the wider issues of the environment that affect engineering economic decision making. Students are expected to participate in class discussion when needed.

4. Assessment Tasks/Activities (ATs)

(ATs are designed to assess how well the students achieve the CILOs.)

Assessment Tasks/Activities	CILO No.						Weighting	Remarks
	1	2	3	4	5			
Continuous Assessment: 65 %								
Group Project		✓		✓	✓		40%	
Individual Coursework	✓	✓	✓	✓			25%	
Examination: 35 % (duration: 2 hours)								
Examination	✓	✓	✓	✓	✓		35%	
							100%	

5. Assessment Rubrics

(Grading of student achievements is based on student performance in assessment tasks/activities with the following rubrics.)

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Marginal (D)	Failure (F)
1. Group Project	40%	High	Significant	Moderate	Basic	Not even reaching marginal levels
2. Individual Coursework	25%	High	Significant	Moderate	Basic	Not even reaching marginal levels
3. Examination	35%	High	Significant	Moderate	Basic	Not even reaching marginal levels

Part III Other Information (more details can be provided separately in the teaching plan)

1. Keyword Syllabus

(An indication of the key topics of the course.)

- Margin-based Classification: Regularization Forms; Functional Margin; Functional Entropy; Empirical Processes; Large Deviation Inequality
- Ensemble Methods: Random Forest; Bagging; Adaptive Boosting; Gradient Boosting; Theory of PAC; Additive Modelling; Functional Gradients
- Unsupervised Learning: K-means; Hierarchical Clustering; Spectral Clustering; Model Selection Consistency; Sufficient Dimension Reduction; Sliced Inverse Regression
- Sparse Learning: High-dimensional Data; Ridge Regression; Lasso Regression; Convex and Non-convex Regularization; Diverging Dimension; Variable Selection Consistency
- Neural Networks: Radial Basis Function Networks; Deep Neural Networks; Convolutional Neural Networks; Stacked Denoising Autoencoders; Generative Adversarial Networks; Function Approximation; Error Analysis

2. Reading List

2.1 Compulsory Readings

(Compulsory readings can include books, book chapters, or journal/magazine articles. There are also collections of e-books, e-journals available from the CityU Library.)

1.	Statistical Inference by George Casella and Roger L. Burger
2.	The Elements of Statistical Learning by Hastie, Tibshirani, and Friedman, Springer
3.	Lecture Notes

2.2 Additional Readings

(Additional references for students to learn to expand their knowledge about the subject.)

NIL