City University of Hong Kong
Course Syllabus

offered by College of Business
with effect from Semester A 2018/19

Part I Course Overview

Course Title:	**Machine Learning for Business Research**
Course Code:	**FB8918**
Course Duration:	**1 semester**
Credit Units:	**3**
Level:	**R8**
Proposed Area: *(for GE courses only)*	☐ Arts and Humanities ☐ Study of Societies, Social and Business Organisations ☐ Science and Technology
Medium of Instruction:	**English**
Medium of Assessment:	**English**
Prerequisites: *(Course Code and Title)*	**Nil**
Precursors: *(Course Code and Title)*	**Students must have taken at least one statistics course at undergraduate/postgraduate level**
Equivalent Courses: *(Course Code and Title)*	**Nil**
Exclusive Courses: *(Course Code and Title)*	**Nil**

Part II Course Details

1. Abstract

(A 150-word description about the course)

Machine learning stands the core for many business models nowadays. This course aims to teach doctoral students in College of Business machine learning models and tools and enable them to conduct related business research. The course will cover supervised learning in depth, including regression, classification, regularization, tree-based methods, ensemble methods etc., and will also introduce the basic concepts and tools of unsupervised learning, including clustering and principle component analysis, etc. This course focuses on practical training using business data, including marketing and financial market data, as well as unstructured text data in news media.

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.	Formulate business research problems using statistical machine learning models.		✓	✓	✓
2.	Analyze and interpret the results of statistical machine learning analyses.		✓	✓	✓
3.	Demonstrate competence in using software packages to analyze business data with statistical machine learning tools.			✓	✓
4.	Communicate and present the results effectively in written, oral and electronic formats.			✓	✓
** If weighting is assigned to CILOs, they should add up to 100%.*		100%			

[#] Please specify the alignment of CILOs to the Gateway Education Programme Intended Learning outcomes (PILOs) in Section A of Annex.

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			
Lecture	Concepts and specific subject knowledge are explained	✓	✓					2.0
Class Discussion	Research problems and research papers are given in class for discussion. Students will be asked to explore possible solutions to these problems and evaluate methods employed in the papers.		✓	✓				0.5
Computer Laboratory Sessions	Computer laboratory sessions provide demonstration and hand-on experience of using statistical packages to analyse datasets. Students have to formulate the research problems into a statistics model and analyse the data with the support of the statistical packages.	✓	✓	✓				0.5
Homework and Project	Research problems with data are assigned to the class. Students, who can work as group, have to integrate the techniques learned in the course to analyse the dataset... Interpretations of the results have to be presented in written or oral format.	✓	✓	✓	✓			N.A.

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				
Continuous Assessment: __50__%								
Group project	✓	✓	✓	✓				
Homework	✓	✓	✓	✓				
In-class participation (computer laboratory sessions)	✓	✓	✓					
Examination: __50__% (duration: 3 hours)								
Examination	✓	✓	✓					
** The weightings should add up to 100%.*							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 and Homework	Evidence of original thinking, organisation, ability to analyse, and grasp of knowledge.	Strong evidence of original thinking; good organization, capacity to analyse and synthesize; superior grasp of subject matter; evidence of extensive knowledge base.	Sufficient evidence of original thinking, some evidence of critical capacity and analytic ability; reasonable understanding of issues; evidence of familiarity with course materials.	Some evidence of original thinking, little evidence of critical capacity and analytic ability; reasonable understanding of course materials.	Little evidence of original thinking, little evidence of critical capacity and analytic ability; reasonable understanding of course materials.	No evidence of familiarity with the subject matter; weakness in critical and analytic skills; limited or irrelevant use of course materials.
2. In-class participation	Understanding of key concepts and definitions, willingness to participate.	Strong evidence of showing understanding of key concepts and definitions; clearly and correctly state most critical points and important contributions of the assigned questions or problems; high participation and excellent presentation skills.	Sufficient evidence of showing understanding of key concepts and definitions; clearly and correctly state some critical points and contributions of the assigned questions or problems; high participation and good presentation skills.	Evidence of showing some understanding of the subject; demonstrate some ability to develop solutions to simple and basic problems in the assigned questions and problems.	State a few critical points and marginal contributions of the assigned questions and problems.	Do not show any participation

3. Examination	Evidence of original thinking, organisation, ability to analyse, and grasp of knowledge.	Strong evidence of original thinking; good organization, capacity to analyse and synthesize; superior grasp of subject matter; evidence of extensive knowledge base.	Sufficient evidence of original thinking; sufficient evidence of critical capacity and analytic ability; reasonable understanding of issues; evidence of familiarity with course materials.	Some evidence of original thinking; some evidence of critical capacity and analytic ability; some understanding of issues; some evidence of familiarity with course materials.	Little evidence of original thinking; little evidence of critical capacity and analytic ability; some understanding of issues; some evidence of familiarity with course content.	Little evidence of familiarity with the subject matter; weakness in critical and analytic skills; limited or irrelevant use of course materials.

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.)

1. Introduction
   Review of basic knowledge on linear algebra, probability theory, and convex optimization. Overview of the concepts of machine learning.

2. Supervised learning
   Regression and classification. Regularization. Generative/discriminative learning. Naive Bayes. Random forest. Support vector machines. Model selection and feature selection. Ensemble methods: bagging, boosting.

3. Unsupervised learning
   K-means clustering. Expectation-maximization. Factor analysis. Principal components analysis (PCA). Independent components analysis (ICA).

4. Project training using Python
   Classification application using marketing data. Prediction problems using financial market data. Natural language processing (NLP) and applications using unstructured textual data from news media.

5. Relevant research paper study from journals including the Accounting Review, Journal of Political Economics, Journal of Financial Economics, Marketing Science, Management Science etc.

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.	Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, An Introduction to Statistical Learning. Springer, 2013.

2.2 Additional Readings

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

1.	Trevor Hastie, Robert Tibshirani and Jerome Friedman, The Elements of Statistical Learning. Springer, 2009
2.	Christopher M. Bishop, Pattern Recognition and Machine Learning. Springer, 2011
3.	Kevin P. Murphy, Machine Learning: A Probabilistic Perspective. The MIT Press, 2012
4	Ian Goodfellow, Yoshua Bengio, Aaron Courville, Deep Learning. The MIT Press, 2016