THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY
Approval of Undergraduate Course

Section 1: Academic Administration [1]

1.1 Catalog

a) Course to be effective from: Academic Year 2021-22 Term Fall

b) Department Code[3]: MATH Subject Area[3]: *MATH* Course Number [4]: 4632

Previous Course Code[5]: MATH4824B (Alternate code: COMP 4901K)

c) Full Title[6] *(max. 100 characters)*: Machine Learning with Structured Data

d) Abbreviated Title[7] *(max. 30 characters)*: ML with Structured Data

e) Course Credits[8]: ☑ Fixed: 3 ○ Range: From ____ To ____

f) Catalog Description[9] *(word limit = 150)*:

This course provides an introduction to statistical machine learning algorithms for structured data such as text sequences, taxonomy trees, relational databases (such as knowledge bases), and graphs (including graph databases such as biomedical graphs and large heterogeneous information networks such as knowledge graphs), and using programming tools such as Python to implement them for real problems. It will use some of the following practical problems such as text and graph classification, statistical relational learning, information extraction, sequence modeling, graph modeling, protein 3D structure prediction, QA system, etc. as illustrations to demonstrate the power of the statistical learning algorithms.

g) Grading Type[10]: ☑ Letter Grades ○ Distinction/Credit/Pass/Fail ○ Pass/ Fail ○ Distinction/Pass/Fail ☐ Others *(please specify)*: ____

h) ☑ Prerequisites[11]:

Course Code / Public Exam	Course Title / Exam Subject and Level / Grade attained
(COMP 2011 OR COMP 2012 OR COMP 2012H) AND	Programming with C++ OR Object-Oriented Programming and Data Structures OR Honors Object-Oriented Programming and Data Structures
(COMP 2711 OR COMP 2711H OR MATH 2343) AND	Discrete Mathematical Tools for Computer Science OR Honors Discrete Mathematical Tools for Computer Science OR Discrete Structures
(MATH 2111 OR MATH 2121 OR MATH 2131)	Matrix Algebra and Applications OR Linear Algebra OR Honors in Linear and Abstract Algebra I

i) ☐ Corequisites[12]:

Course Code	Course Title

j) ☑ Exclusions[13]:

Course Code / Public Exam	Course Title / Exam Subject and Level / Grade attained
COMP 4901K	Machine Learning for Natural Language Processing
MATH 4824B	Machine Learning for Natural Language Processing

k) ☐ Co-listing[14]: ☑ Multi-coding[14]:

Course Code	Course Title
COMP 4222	Machine Learning with Structured Data

l) Other Enrollment Restrictions[15] ◉ No ○ Yes

☐ Instructor's approval required

☐ Restricted to specified student group(s)
(please specify, e.g. year and program of study): _____

☐ Others *(please specify):* _____

m) Medium of Instruction/Materials[16]: ◉ English ○ Others, (Pls specify and provide a justification in Section 1.3): _____

n) Allow course repetition for credit[17]: ◉ No ○ Yes

1.2 Contribution of course to Programs of Study *[Check all appropriate boxes below]*

☑ Major

Program of Study	As		
BEng(COMP) BSc(COSC) BSc(DSCT) BSc(MATH)	☐ Required Course	☑ Elective	☐ Prerequisite

☐ Minor

Program of Study	As		
	☐ Required Course	☐ Elective	☐ Prerequisite

☐ Common Core

☐ Others (pls specify):

Program of Study	As		
	☐ Required Course	☐ Elective	☐ Prerequisite

1.3 Rationale for Introducing this course and other relevant information [18]

The course is an interdisciplinary course which needs both computer science background and mathematics background. The algorithms introduced in this course will enrich both CSE and Math students' knowledge. The spectral graph theory is highly related to Math, which is the foundation of development of graph neural networks. Then the realization and implementation of machine learning algorithms of structured data is highly related to CSE techniques. The students are required to work in small groups for a number of homework assignments. During the course, there will be some projects requires students working as teams to work on some real world problems. It will encourage students from Math and CSE (especially DCST), CPEG, and other departments to register and to work together to bring different background knowledge working on interesting real problems. The students will be merged in one Canvas session so they can collaborate with each other to work on the assignments/projects. This will enable students to form multidisciplinary teams. The course especially fits the DCST program with complementary contents in additional to existing machine learning and optimization courses to deal with more complex data structures. The multi-coded courses will be identical to students enrolled in both course codes. The evaluation, examination, projects, assignments will be identical.

Section 2A: Learning Outcomes and Alignment *(for courses not proposed to be Common Core Courses)*

2.1 Key Course Intended Learning Outcomes *(Should not normally exceed six or eight outcomes)*

Upon completion of this course, students are expected to be able to do the following:

	Course ILOs	**Nature of the learning outcomes** *(A - Knowledge/Content Related; B - Academic Skills/Competencies; C - Others)*
1	Explain the basic principles behind machine learning algorithms for structured data	A
2	Implement programs for structured prediction tasks	B
3	Formulate machine learning solutions to domain problems	B
4	Demonstrate the ability to understand of the complexity of real world problems	B

2.2 Contribution of Learning Outcomes to Programs of Study identified in Section 1.2

(Please also complete Section 4.1)

	Program of study 1: _BEng(COMP) **Program ILOs**	**To be achieved through these course ILOs** *(Write CILO-1, CILO-2, etc.)*
1	PO1. An ability to apply knowledge of computing and mathematics appropriate to the discipline.	CLIO-1
2	PO2. An ability to apply knowledge of a computing specialisation, and domain knowledge appropriate for the computing specialisation to the abstraction and conceptualisation of computing models.	CLIO-2, CLIO-3
3	PO3. An ability to analyze a problem, and identify and define the computing requirements appropriate to its solution.	CLIO-2, CLIO-3
4	PO4. An ability to design, implement and evaluate a computer-based system, process, component, or program to meet desired needs.	CLIO-2, CLIO-3, CLIO-4
5	PO5. An ability to function effectively in teams to accomplish a common goal.	CLIO-4
6	PO6. An understanding of professional, ethical, legal, security and social issues and responsibilities.	
7	PO7. An ability to communicate effectively with a range of audiences	CLOI-4
8	PO8. An ability to analyze the local and global impact of computing on individuals, organizations, and society.	
9	PO9. Recognition of the need for, and an ability to engage in, continuing professional development.	
10	PO10. An ability to use current techniques, skills, and tools necessary for computing practices.	CLOI-4

	Program of study 2: BSc(COSC) **Program ILOs**	**To be achieved through these course ILOs** *(Write CILO-1, CILO-2, etc.)*
1	Explain knowledge, principles and use of IT skills in mathematical and computer sciences at college level. (Knowledge)	CILO-1
2	Evaluate information critically, and make independent judgment by applying principles and methods in mathematical and computer sciences. (Judgment)	CILO-1,CILO-3
3	Apply quantitative, analytic methods and IT skills to execute tasks and solve problems in mathematical and computer sciences. (Execution)	CILO-2
4	Work independently and collaborate effectively in a team. (Interpersonal Skill and Leadership)	CILO-3,CILO-4
5	Communicate effectively, both in oral and written forms, about mathematical knowledge to audience. (Communication)	CILO-3,CILO-4
6	Self-evaluate their own learning progress, and develop motivation and skills for lifelong learning. (Self-reflection)	

7	Recognize the importance of complying with ethics of science and academic integrity. (Ethical Practice)	
8	Show appreciation of mathematical and computer sciences and its interface with human activities, and arouse audience's interest in the beauty, logic and precision of mathematical and computer sciences. (Appreciation)	CILO-1
9	View issues in mathematical sciences with reference to the practices of the international science community. (International Outlook)	

	Program of study 3: BSc(DSCT) **Program ILOs**	**To be achieved through these course ILOs** *(Write CILO-1, CILO-2, etc.)*
1	The ability to understand data problems arising in the areas of commerce and industry etc.	CILO-3,CILO-4
2	The ability to model data problems using different mathematical tools.	CILO-1
3	The ability to design and implement efficient algorithms to solve different mathematical models for data problems.	CILO-2
4	The ability to interpret the results provided by different algorithms and apply them to the data problems to gain meaningful insights or offer predictions.	CILO-3,CILO-4

	Program of study 4: BSc(MATH-AM) **Program ILOs**	**To be achieved through these course ILOs** *(Write CILO-1, CILO-2, etc.)*
1	Explain knowledge, principles and use of quantitative techniques in mathematical sciences at college level. (Knowledge)	CILO-1
2	Model real-world problems and information mathematically, and make independent judgment by applying structural and analytical approaches. (Judgment)	CILO-1
3	Apply logical, analytic, and highly numerate methods to execute tasks and solve real-world mathematical problems. (Execution)	CILO-1
4	Work independently and collaborate effectively in a team. (Interpersonal Skill and Leadership)	CILO-2,CILO-3
5	Communicate effectively, both in oral and written forms, about mathematical knowledge to audience. (Communication)	CILO-4
6	Self-evaluate their own learning progress, and develop motivation and skills for lifelong learning. (Self-reflection)	CILO-4
7	Recognize the importance of complying with ethics of science and academic integrity. (Ethical Practice)	
8	Show appreciation of mathematical sciences and its interface with human activities, and arouse audience's interest in the beauty, logic and precision of mathematical sciences. (Appreciation)	CILO-1
9	View issues in mathematical sciences with reference to the practices of the international science community. (International Outlook)	

Section 2B: Additional Information[2] *(for courses not proposed to be Common Core Courses)*

2.3 Planned Teaching & Learning Arrangement

Teaching & Learning Arrangement		Weekly Scheduled Hours/ Estimated Weekly Learning Hours	Indicate which course ILOs this activity serves to achieve *(Write CILO-1, CILO-2, etc.)*	Additional Information *(optional)*
Face-to face activities	☑ Lecture*	3 hours	CILO-1, CILO-2, CILO-3, CILO-4	
	☑ Tutorial*	1-hour	CILO-1, CILO-2, CILO-3, CILO-4	
	☐ Seminar/Small-class*			
	☐ Laboratory*			
	*Does the above scheduled component(s) involve structured active learning activities? [19] ◉ No ○ Yes *If yes, please specify for each scheduled component, the percentage and the type of active learning involved in the "Additional Information" column.*			
	☐ Others (e.g. fieldtrip, visit, etc.), *pls specify:* _____			
Online activities	☐ Online lecture videos			
	☐ Other online learning tasks, *pls specify:* _____			
The total learning hours of the course# is equivalent to 120 **hours** [8] *# including both scheduled instructional hours and hours for self-study activities & assessment*				

- *For course adopting a pedagogic approach other than lecture, tutorial and laboratory, please indicate the pedagogy used:*

○ Blended learning [20] ○ Pure online delivery [21]

○ Experiential learning [22] ○ Others, *pls specify:*_____

2.4 Planned Assessment Weightings

Assessment Task	Proportion of Final Grade (%)	Indicate which course ILOs this task is to assess *(Write CILO-1, CILO-2, etc.)*	Additional Information *(optional)*
☐ In-class test			
☐ Mid-term test			
☑ Final exam	40%	CILO-1, CILO-3	
☑ Assignments	30%	CILO-1, CILO-2, CILO-3	
☑ Final Project	20%	CILO-2, CILO-3, CILO-4	
☑ Presentation	10%	CILO-3, CILO-4	
☐ Learning portfolio			
☐ Course participation			
☐ Peer evaluation			
☐ Others (e.g. proctored online exam, etc.), *pls specify:* _____			

2.5 Course Duration

☑ 1 term ○ 2 terms ○ Others, *pls specify:* _____

***2.6 Planned Frequency of Offerings** [Check all appropriate boxes]:*

☐ Every Fall ☐ Every Winter

☐ Every Spring ☐ Every Summer

☐ No fixed pattern

☑ Other *(pls specify):* This course (COMP4222/MATH4632) will be taught every two years. The other PG co-listed course (COMP5222/MATH5471) will be taught with similar purpose.

2.7 Course outline attached ○ No ☑ Yes

- **Internationalization:**

Internationalization in a course refers to course content and/or pedagogic approaches which incorporate an intercultural and international perspective. Examples may include:

- Collaboration with overseas institutions to develop and adopt international course content, or to arrange international field trip

- Insertion of international theme as part of the course

- Integrating the course content with international material as examples or case studies

- Elements to provide global diversified perspectives and/or practices around the world

Please briefly list or summarize any component(s) in the course that contributes to internationalizing the curriculum:

2.8 Resources

Request extra resources for teaching this course? ○ No ☑ Yes

Textbook / Reference Books

- Jurafsky and Martin (2008), Speech and Language Processing, 2nd edition.
- Noah Smith (2011), Linguistic structure prediction, Online.
- Lise Getoor and Ben Taskar (2007). Introduction to Statistical Relational Learning. The MIT Press.
- Pedro Domingos and Daniel Lowd, Markov Logic: An Interface Layer for AI, Morgan & Claypool, 2008.

Course Outline of COMP4222 (multi-coding with MATH4632)

Week	Topics	Briefly outline what this topic will cover *(Include reading assignments if available)*	Indicate which course ILOs this topic is related to *(Write CILO-1, CILO-2, etc.)*
1	Introduction	Introduction to the course and context of the content.	CILO-1
2	Structured perceptron and its generalizations with global optimization methods	Introduction to structure prediction problems and the basic algorithms, Relational Markov networks and conditional random fields	CILO-1
3	Graph based semi-supervised learning	Spectral graph theory, graph Laplacian	CILO-1
4	Introduction to deep learning	Introduction basic deep learning concepts for structured data, e.g., CNN, RNN on node classification, link prediction over sequences, trees, and graphs	CILO-1
5	Network embedding	Deepwalk, node2vec, heterogeneous information network embeddings, etc.	CILO-1
6	Deep sets	Generalize deep learning algorithms to set data, Transformer Networks	CILO-1
7	Graph neural networks	General graph neural networks: Graph CNN, GraphSage, Message Passing Networks	CILO-1
8	Graph isomorphism and subgraph isomorphisms	Graph Isomorphism Networks and applications such as summary statistics, counting, other NP hard problems	CILO-1
9	Deep graph generation	Generative models for graphs	CILO-1
10	Application 1: Knowledge graph base QA System	QA system using existing knowledge graphs	CILO-3, CILO4
11	Application 2: Protein 3D structure prediction	AlphaFold and others in biomedical data	CILO-3, CILO4
12	Student project presentations	Knowledge sharing	CILO-2, CILO-3, CILO4
13	Student project presentations	Knowledge sharing	CILO-2, CILO-3, CILO4

Section 3: Learning Outcomes and Alignment *(for Common Core Course)*

3.1 Course Objectives: *Please outline what this course aims to achieve*

Alignment with Common Core program goals (Details here): Check the appropriate box(es) below to indicate which Common Core goal(s) this course aims to achieve, and explain briefly how this course would help to achieve the selected Common Core goal(s).

Section 4: Development, Concurrence and Approval

4.1 Contribution to the Program Learning Outcomes

The course is confirmed by the following Major/Minor program department(s)/unit(s) as indicated in Section 1.2 that it would contribute appropriately to overall program learning outcomes.

Department/Program unit	*Position*	*Name*	*Date*
Dept of Computer Science and Engineering	UG Coordinator	Dr Qiong LUO	14-Jan-21
Dept of Mathematics	Program Director	Prof Mo MU	18-Jan-21
Dept of Mathematics	UG Coordinator	Dr Tsz Kin LAM	19-Jan-21

4.2 Approvals

Recommendation from offering department(s) and School(s)/IPO

Offering Department/Program Unit	*Position*	*Name*	*Date*
Dept of Computer Science and Engineering	UG Coordinator	Dr Qiong LUO	14-Jan-21
Dept of Mathematics	UG Coordinator	Dr Tsz Kin LAM	19-Jan-21

Recommending School/IPO	*Position*	*Name*	*Date*
School of Engineering	Associate Dean	Prof Philip MOK	19-Feb-21
School of Science	Associate Dean	Prof Pak Wo LEUNG	19-Feb-21

Concurrence from other Schools or departments/units

School/Dept/Program Unit	*Position*	*Name*	*Date*
Interdisciplinary Programs Office	UG Coordinator	Prof Betty LIN	15-Jan-21