Applicable to 42477-1 Students
(2019/20 Intake Cohort)

The Hong Kong Polytechnic University
Department of Electronic and Information Engineering

Minor Changes to the BSc (Hons) in Internet and Multimedia Technologies (BSc in IMT) (42477) Programme

Background

The Department reviews the programme regularly to ensure the subjects of the curriculum match with the education background and needs of the students, align with the programme aims, objectives and intended learning outcomes, cater to the demand for talents from the industries, and stay abreast of the development of the technology. To align with the development trend of many electronic and information engineering as well as computer science academic programmes around the world where training on Artificial Intelligence (AI) is introduced into the curricula, the Department has implemented a number of relevant minor changes to the BSc (Hons) in Internet and Multimedia Technologies (BSc in IMT) (42477) programme.

1. Removing "EIE1002 Electronics Science" from and Adding a New Subject "EIE1003 Foundations of Data Science" to the 4-year Curriculum

The first change is to remove the existing Level 1 compulsory subject, "EIE1002 Electronics Science" from the 4-year curriculum and add a new compulsory subject, "EIE1003 Foundations of Data Science" (Appendix I), to the 4-year curriculum as a replacement.

The Department has considered that being able to discover useful knowledge and information from a large amount of data is very critical to industry, business and government. The new subject "EIE1003 Foundations of Data Science" aims to provide students with the fundamental concepts of data science and the basic technologies for data analytics. It provides hands-on experiences in data analytics and case studies in applications of data science in engineering, social science, healthcare, business and government. It also prepares students with the right mentality towards data and the ability to leverage data for decision-making. Students admitted to the programme through the normal year 1 entry are expected to take EIE1003 in Year 1 Semester 2.

Such changes will take effect from 2019/20 Semester 2 and be applicable to 2019/20 intake cohort of the normal year 1 entry and beyond.

2. Adding an elective subject "EIE4121 Machine Learning for Cyber-security" to the 4-year and Senior Year Curricula

Machine learning techniques have become popular in many different fields. In recent years, machine learning techniques have been applied to cyber-security tasks. There is a growing trend of using these techniques to cyber-security tasks such as application and endpoint security. Currently, the BSc in IMT programme does not offer any subject that covers machine learning techniques in cyber-security. It is beneficial to students to learn these techniques and develop skills to apply these techniques to solve various problems in cyber-security.

Besides, during the student feedback collection exercise in 2018/19 Semester 2, the Department sought students' views on the area(s) of subject(s) they preferred if the Department had resources to offer new technical elective subjects for the undergraduate degree programmes it offered. Most students who had given comments preferred new subjects in the areas of Artificial Intelligence, Big Data and Machine Learning. The Department considered that a new subject in the field of Machine Learning could be offered to cater to students' interests and respond to the Departmental Review panel members' suggestions.

The Department therefore introduces a new 3-credit subject "EIE4121 Machine Learning in Cyber-security" (Appendix II) to the 4-year and senior year curricula of the BSc in IMT programme as a technical elective subject. This subject intends to introduce concepts about machine learning techniques in cyber-security to students and develop students' skills of using recent techniques for solving practical problems in cyber-security. Topics such as machines learning techniques, machine learning development environments, malware analysis, phishing detection and anomaly detection will be covered in this subject.

Adding EIE4121 to the 4-year and senior year curricula provides BSc in IMT students with more choices of elective subjects.

The adding of "EIE4121 Machine Learning in Cyber-security" to the 4-year and senior year curricula of the BSc in IMT programme as a technical elective will take effect from 2019/20 Semester 2 and be applicable to all students of the BSc in IMT programme.

Subject Description Form

Subject Code	EIE1003
Subject Title	Foundations of Data Science
Credit Value	3
Level	1
Pre-requisite/ Co-requisite/ Exclusion	Nil
Objectives	Being able to discover useful knowledge and information from a large amount of data is very critical to industry, business and government. This subject aims to provide students the fundamental concepts of data science and the basic technologies for data analytics. It provides hands-on experiences in data analytics and case studies in applications of data science in engineering, social science, healthcare, business and government. It also prepares students with the right mentality towards data and the ability to leverage data for decision-making.
Intended Subject Learning Outcomes	**Upon completion of the subject, students will be able to:** Category A: Professional/academic knowledge and skills 1. Understand the basic concepts and technologies of data science. 2. Acquire the basic technical know-how on data analytics. Category B: Attributes for all-roundedness 3. Understand the data-driven process for problem solving. 4. Demonstrate how to harness and process data for decision-making.
Contribution of the Subject to the Attainment of the Programme Outcomes	Category A: Professional/academic knowledge and skills • Programme Outcome 2, 5: This subject contributes to these programme outcomes through the teaching of fundamental data science concepts and practical training in lab sessions. Category B: Attributes for all-roundedness • Programme Outcome 6, 8: This subject contributes to the programme outcome through group projects and group presentation. • Programme Outcome 7, 9: This subject contributes to the programme outcome through the teaching of professional ethic in handling data and the impact of data science on society, business and government.
Subject Synopsis/ Indicative Syllabus	1. Introduction to Data Science • Data science vs. big data vs. data analytics • Benefits of data science • Skill sets required • Privacy, security and ethics • Example applications and case studies 2. Technologies for Data Science • Basic concepts in summary statistics • Graphs and plots for data analytics, e.g., box plots, scatter plots, histograms, run charts, etc. • Example case studies of exploratory data analytics for data science • Fundamental of machine learning for data science • Cloud technologies 3. Tools for Data Science • Data cleaning, e.g., OpenRefine • Machine learning tools, e.g., Microsoft ML Studio, Weka • Data visualization tools, e.g., Google Chart, Tableau

<table>
<tr><td></td><td>4. Applications with Case Studies
Recommendation systems
Spam filtering
Stock prediction
Social networks
Sentiment analysis</td></tr>
<tr><td>Teaching/Learning Methodology</td><td>Lectures: The subject matters will be delivered through lectures (both in-person and online ones). Students will be engaged in the lectures through Q&A, discussions and specially designed classroom activities. Practitioners and software vendors will be invited to give guest lectures.

Tutorials and Workshops: Students will work on data analytics projects using software tools. Students will start from small and easy projects in the first half of the subject. In the second half, students will work on a more realistic project that solves real-world problems, using the knowledge and know-hows that they have learnt from the small projects.

Assignment: Students will need to do a group-based mini-project on data science.</td></tr>
<tr><td>Assessment Methods in Alignment with Intended Subject Learning Outcomes</td><td>
<table>
<tr><th rowspan="2">Specific Assessment Methods/Tasks</th><th rowspan="2">% Weighting</th><th colspan="4">Intended Subject Learning Outcomes to be Assessed (Please tick as appropriate)</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>4</th></tr>
<tr><td>1. Continuous Assessment (total: 100%)</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>• Mini-project (proposal, report and presentation)</td><td>40%</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td></tr>
<tr><td>• Tests and Quizzes</td><td>30%</td><td>✓</td><td>✓</td><td></td><td></td></tr>
<tr><td>• Laboratory exercises</td><td>30%</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td></tr>
<tr><td>Total</td><td>100%</td><td colspan="4"></td></tr>
</table>
Explanation of the appropriateness of the assessment methods in assessing the intended learning outcomes:

Laboratory exercises and mini-project will require students to apply what they have learnt to solve problems. There will be open-ended questions that allow students to exercise their creativity in solution design.

Tests and Quizzes assess students' achievement of the learning outcomes in a more formal manner.

Mini-project is group-based and weights 40% of the whole assessment. Among the 40% weight, 10% is for proposal, 20% is for final report, and 10% is for presentation (in the form of a 10-minutes video). Proposal and report (30% in total) are evaluated based on group, while presentation (10%) is evaluated individually. Each group member will present the part he/she is responsible for in the mini-project. The mini-project will make use of publicly available tools such as Microsoft Azure Machine Learning Studio so that requirements on programming knowledge is kept to a minimum, i.e., no programming background is assumed. Students will perform drag and drop of data sources, machine learning models, analytic methods, and evaluation methods from the tool to solve data science problems. Enthusiastic students could use the cloud-based API to perform more complex tasks.

Tests and Quizzes weight 30% and they are individual assessments.

Laboratory exercises weight 30%. Half of them (15%) are individual assessments and half of them (15%) are group-based assessments.</td></tr>
</table>

	Overall, 55% of the assessment is individual assessment and 45% is group-based assessment.	
Student Study Effort Expected	**Class contact (time-tabled):**	
	• Lectures (In-person and online)	22 Hours
	• Tutorial/Laboratory/Practice Classes	17 Hours
	Other student study effort:	
	• Lecture: preview/review of notes; homework/assignment; preparation for test/quizzes	30 Hours
	• Tutorial/Laboratory/Practice Classes: preview of materials, revision and/or reports writing	36 Hours
	Total student study effort:	**105 Hours**
Reading List and References	**Reference Materials:** 1. L. B. Cao, “Data Science Thinking: The Next Scientific, Technological and Economic Revolution”, Springer, 2018. 2. V. Mayer-Scgibberger abd K. Cukier, “Big Data: A Revolution That Will Transform How We Live, Work and Think”, John Murray, 2013. 3. M. Manoochehri, “Data Just Right: Introduction to Large-Scale Data & Analytics, Addison-Wesley, 2014. 4. S. Smolan, B. Medsker and Joel McHale “The Human Face of Big Data”, PBS Distribution, 2016 (DVD Video). 5. P. Kromer and R. Jurney, “Big Data for Chimps”, O’Reilly, 2016. 6. T. Ojeda et al. Practical data science cookbook. Packt Publishing Ltd, 2014. 7. Computational and Inferential Thinking: The Foundations of Data Science. Online textbook, https://www.inferentialthinking.com/ 8. A list of web references for each topic.	
Last Updated	June 2019	
Prepared by	Man-Wai MAK	

Subject Description Form

Subject Code	EIE4121
Subject Title	Machine Learning in Cyber-security
Credit Value	3
Level	4
Pre-requisite	Nil
Co-requisite/ Exclusion	Nil
Objectives	1. To introduce concepts about machine learning techniques in cyber-security 2. To develop skills of using recent techniques for solving practical problems in cyber-security
Intended Learning Outcomes	**Upon completion of the subject, students will be able to:** Category A: Professional/academic knowledge and skills 1. Understand different machine learning techniques 2. Use different techniques for solving problems in cyber security Category B: Attributes for all-roundedness 3. Present ideas and findings effectively
Subject Synopsis/ Indicative Syllabus	**Syllabus:** 1. Machine learning techniques Introduction to machine learning; Basic concepts and classification; Supervised learning and unsupervised learning; classification; clustering; Neural Networks; Support vector machines; Dimensionality reduction; Deep learning 2. Machine learning development environments Software tools for implementing machine learning techniques; Generalization performance; Issues of over-fitting. 3. Malware Analysis Introduction to malware analysis; Types of malware analysis; static analysis, dynamic analysis; Behavioral vs code analysis; Use of machine learning techniques for malware detection such as K-Means, support vector machines, convolutional neural networks. 4. Phishing detection Introduction to phishing detection; Analysis of email/websites/message features for phishing characterization; Use of techniques such as logistic regression and decision tree for phishing detection. 5. Anomaly Detection Introduction to the anomaly definition; overview of anomaly detection techniques; static rules technique; use of machine learning techniques such as autoencoder for anomaly detection. **Laboratory Experiments:** Practical Works: 1. Evaluation of machine learning techniques in malware detection 2. Evaluation of machine learning techniques in phishing detection Forensic analysis of digital evidence.

Teaching/Learning Methodology

Teaching and Learning Method	Intended Subject Learning Outcome	Remarks
Lectures	1, 2	Fundamental principles and key concepts of the subject are delivered to students.
Tutorials	1, 2	Supplementary to lectures; Students will be able to clarify concepts and to have a deeper understanding of the lecture material; Problems and application examples are given and discussed.
Laboratory sessions	2, 3	Students will evaluate different kinds of machine learning techniques.
Mini-project	1, 2, 3	Students are required to study the use of machine learning techniques in cyber-security application. Students will need to submit a written report and make a presentation.

Assessment Methods in Alignment with Intended Learning Outcomes

Specific Assessment Methods/Tasks	% Weighting	Intended Subject Learning Outcomes to be Assessed (Please tick as appropriate)		
		1	2	3
1. Continuous Assessment (total 50%)				
• Tests	17%	√	√	
• Short quizzes	10%	√	√	
• Laboratory sessions	5%		√	√
• Mini-project	18%		√	√
2. Examination	50%	√	√	
Total	100%			

The continuous assessment consists of tests, short quizzes, laboratory exercises and a mini-project.

Explanation of the appropriateness of the assessment methods in assessing the intended learning outcomes:

Specific Assessment Methods/Tasks	Remark
Short quizzes	These can measure students' understanding of the theories and concepts as well as their comprehension of subject materials.
Tests and examination	end-of chapter type problems used to evaluate students' ability in applying concepts and skills learnt in the classroom; students need to think critically in order to come with a solution for a problem.
Laboratory sessions, mini-project	oral examination will be conducted to evaluate student's technical knowledge and communication skills.

Student Study Effort Expected	**Class contact (time-tabled):**	
	• Lecture	24 Hours
	• Tutorial/Laboratory/Practice Classes	15 Hours
	Other student study effort:	
	• Lecture: preview/review of notes; homework/assignment; preparation for test/quizzes/examination	26 Hours
	• Tutorial/Laboratory/Practice Classes: preview of materials, revision and/or reports writing	40 Hours
	Total student study effort:	**105 Hours**
Reading List and References	1. Mark Stamp, Introduction to Machine Learning with Applications in Information Security, Chapman and Hall/CRC, 2017. 2. Chiheb Chebbi, Mastering Machine Learning for Penetration Testing, Packt Publishing Ltd, 2018. 3. Sumeet Dua and Xian Du, Data Mining and Machine Learning in Cybersecurity, Auerbach Publications, 2011. 4. Monnappa K A, Learning Malware Analysis, Packt Publishing Ltd, 2018. 5. Dipanjan Sarkar, Raghav Bali and Tushar Sharma, Practical Machine Learning with Python, Apress, 2018.	
Last Updated	1 Feb 2019	
Prepared by	Bonnie Law	