REGULATIONS FOR THE DEGREE OF BACHELOR OF ENGINEERING IN DATA SCIENCE AND ENGINEERING [BEng(DS&E)]

These regulations apply to students admitted to the four-year BEng in Data Science and Engineering curriculum in the academic year 2022-23 and thereafter.
(See also General Regulations and Regulations for First Degree Curricula)

Admission to the Degree

DS&E 1 To be eligible for admission to the degree of Bachelor of Engineering in Data Science and Engineering [BEng(DS&E)], a candidate shall

(a) comply with the General Regulations;
(b) comply with the Regulations for First Degree Curricula; and
(c) satisfy all the requirements of the curriculum in accordance with these regulations and the syllabuses.

Period of Study

DS&E 2 The curriculum for the BEng(DS&E) degree shall normally require eight semesters of full-time study, extending over not fewer than four academic years, and shall include any assessment to be held during and/or at the end of each semester. Candidates shall not in any case be permitted to extend their studies beyond the maximum period of registration of six academic years, unless otherwise permitted by the Board of the Faculty of Engineering.

Curriculum Requirements and Progression in Curriculum

DS&E 3

(a) Candidates shall satisfy the requirements prescribed in UG5 of the Regulations for First Degree Curricula;
(b) Candidates shall take not fewer than 240 credits of courses, in the manner specified in these regulations and syllabuses; candidates are also required to pass all courses as specified in the syllabuses;
(c) Candidates shall normally select not fewer than 24 credits nor more than 30 credits of courses in any one semester (except the summer semester), unless otherwise permitted or required by the Board of the Faculty, or except in the last semester of study when the number of outstanding credits required to complete the curriculum requirements is fewer than 24 credits;
(d) Candidates may, of their own volition, take additional credits not exceeding 6 credits in each semester, and/or further credits during the summer semester, accumulating up to a maximum of 72 credits in one academic year. Candidates may, with the approval of the Board of the Faculty, exceed the annual study load of 72 credits in a given academic year provided that the total number of credits taken does not exceed the maximum curriculum study load of 288 credits for the normative period of study as specified in DS&E 2, save as provided for under DS&E 3(e);
(e) Where candidates are required to make up for failed credits, the Board of the Faculty may give permissions for candidates to exceed the annual study load of 72 credits provided that the total number of credits taken does not exceed the maximum curriculum study load of 432 credits for the maximum period of registration specified in DS&E 2.

DS&E 4 Candidates with unsatisfactory academic progress may be required by the Board of the Faculty to take a reduced study load.

Selection of Courses

DS&E 5 Candidates shall select their courses in accordance with these regulations and the guidelines specified in the syllabuses before the beginning of each semester. Changes to the selection of courses may be made only during the add/drop period of the semester in which the course begins, and such changes shall not be reflected in the transcript of the candidate. Requests for changes after the designated add/drop period of the semester shall be subject to the approval of the Board of the Faculty. Withdrawal from courses beyond the designated add/drop period will be subject to the approval of the Board of the Faculty.

Assessment and Grades

DS&E 6 Candidates shall be assessed for each of the courses for which they have registered, and assessment may be conducted in any one or any combination of the following manners: written examinations or tests, continuous assessment, laboratory work, field work, project reports, or in any other manner as specified in the syllabuses. Only passed courses will earn credits. Grades shall be awarded in accordance with UG8 of the Regulations for First Degree Curricula.

DS&E 7 Written examinations or tests shall normally be held at the end of each semester unless otherwise specified in the syllabuses.

DS&E 8 Candidates are required to make up for failed courses in the following manner:

a) undergoing re-assessment/re-examination in the failed course to be held no later than the end of the following semester (not including the summer semester); or
b) re-submitting failed coursework, without having to repeat the same course of instruction; or
c) repeating the failed course by undergoing instruction and satisfying the assessments; or
d) for elective courses, taking another course in lieu and satisfying the assessment requirements.

DS&E 9 Candidates shall not be permitted to repeat a course for which they have received a D grade or above for the purpose of upgrading.

DS&E 10 There shall be no appeal against the results of examinations and all other forms of assessment.

DS&E 11 Unless otherwise permitted by the Board of the Faculty, a candidate will be recommended for discontinuation of his/her studies if

(a) he/she fails to complete successfully 36 or more credits in two consecutive semesters (not including the summer semester), except where they are not required to take such a number of credits in the two given semesters; or
(b) he/she fails to achieve an average Semester GPA of 1.0 or higher for two consecutive semesters (not including the summer semester); or
(c) he/she has exceeded the maximum period of registration specified in DS&E 2.

Absence from Examination

DS&E 12 Candidates who are unable, because of illness, to be present at the written examination of any course may apply for permission to present themselves at a supplementary examination of the same course to be held before the beginning of the First Semester of the following academic year. Any such application shall be made on the form prescribed within seven calendar days of the first day of the candidate's absence from any examination. Any supplementary examination shall be part of that academic year's examinations, and the provisions made in the regulations for failure at the first attempt shall apply accordingly.

Advanced Standing

DS&E 13 Advanced standing may be granted to candidates in recognition of studies completed successfully before admission to the University in accordance with UG2 of the Regulations for First Degree Curricula. The amount of credits to be granted for advanced standing shall be determined by the Board of the Faculty, in accordance with the following principles:

(a) a minimum of four semesters of study at this University shall be required before a candidate is considered for the award of the degree; and
(b) at least half of the credits requirements of the degree curriculum shall be accumulated through study at this University, or from transfer of credits for courses completed at other institutions in accordance with UG4(d) of the Regulations for First Degree Curricula.

Credits granted for advanced standing shall not normally be included in the calculation of the GPA unless permitted by the Board of the Faculty but will be recorded on the transcript of the candidate.

Degree Classification

DS&E 14 To be eligible for the award of the BEng(DS&E) degree, candidates shall have:

(a) achieved a Graduation GPA of 1.00 or above
(b) satisfied all the requirements in UG5 of the Regulations for First Degree Curricula;
(c) passed not fewer than 240 credits, comprising
    i) introductory courses, including Engineering Core courses;
    ii) advanced courses;
    iii) capstone experience;
    iv) a Chinese language enhancement course[1];

---

[1] Students are required to successfully complete the 6-credit Faculty-specific Chinese language enhancement course, except for:

(a) Putonghua-speaking students who should take CUND9002 (Practical Chinese and Hong Kong Society) or CUND9003 (Cantonese for Non-Cantonese Speaking Students); and
(b) students who have not studied Chinese language during their secondary education or who have not attained the requisite level of competence in the Chinese language to take the Chinese language enhancement course should seek approval from the Board of the Faculty of Engineering for exemption from the Chinese language requirement, and
    (i) take a 6-credit Cantonese or Putonghua language course offered by the School of Chinese especially for international and exchange students; OR
    (ii) take an elective course in lieu.

v) two English language enhancement courses, including Core University English[2] and English in the Discipline;
vi) Common Core courses;
vii) all required courses as prescribed in respective syllabuses;
viii) elective courses; and
ix) non-credit bearing courses as required by the University.

DS&E 15 The degree of Bachelor of Engineering in Data Science and Engineering shall be awarded in five divisions: First Class Honours, Second Class Honours Division One, Second Class Honours Division Two, Third Class Honours, and Pass. The classification of honours shall be determined by the Board of Examiners for the degree of BEng in accordance with the following Graduation GPA (GGPA) scores, with all courses taken (including failed courses) carrying weightings which are proportionate to their credit values [3]:

Class of honours	***GGPA range***
First Class	3.60 – 4.30
Second Class	(2.40 – 3.59)
Division One	3.00 – 3.59
Division Two	2.40 – 2.99
Third Class	1.70 – 2.39
Pass	1.00 – 1.69

DS&E 16 Honours classification may not be determined solely on the basis of a candidate's Graduation GPA and the Board of Examiners for the BEng degree may, at its absolute discretion and with justification, award a higher class of honours to a candidate deemed to have demonstrated meritorious academic achievement but whose Graduation GPA falls below the range stipulated in DS&E 15 of the higher classification by not more than 0.1 Grade Point.

DS&E 17 A list of candidates who have successfully completed all degree requirements shall be posted on Faculty notice boards.

[2] Candidates who have achieved Level 5 or above in English Language in the Hong Kong Diploma of Secondary Education Examination, or equivalent, are exempted from this requirement, and Core University English is optional. Those who do not take this course should take an elective course in lieu, see Regulation UG6.

[3] For students in the 2017-18 intake and thereafter who have successfully completed six Common Core courses, the calculation of Graduation GPA is subject to the proviso that either five Common Core courses with the highest grades (covering all four Areas of Inquiry), or all six courses will be counted towards Graduation GPA, depending on which generates the higher Graduation GPA.

SYLLABUS FOR THE DEGREE OF BACHELOR OF ENGINEERING IN DATA SCIENCE AND ENGINEERING [BEng(DS&E)] (Subject to Approval)

The syllabus applies to students admitted in the academic year 2022-23 and thereafter under the four-year curriculum.

Definition and Terminology

Each course offered shall be classified as either introductory level course or advanced level course.

A Discipline Core course is a compulsory course which a candidate must pass in the manner provided for in the Regulations.

A Discipline Elective course refers to any technical course offered for the fulfillment of the curriculum requirements of the degree of BEng in Data Science and Engineering that are not classified as discipline core course.

Curriculum

The curriculum comprises 240 credits of courses as follows:

Engineering Core Courses
Students are required to complete at least 24 credits of Engineering Core Courses.

Discipline Core Courses
Students are required to complete all discipline core courses (48 credits), comprising 30 credits of introductory core courses and 18 credits of advanced core courses.

Discipline Elective Courses
Students are required to complete at least 30 credits of discipline elective courses offered for the curriculum.

Elective Courses
Students are required to complete 72 credits of elective course(s) offered by any department, except Common Core Courses.

University Requirements
Students are required to complete:

a) 12 credits in English language enhancement, including 6 credits in "CAES1000 Core University English" and 6 credits in "CAES9542 Technical English for Computer Science";
b) 6 credits in Chinese language enhancement course "CENG9001 Practical Chinese for Engineering Students";
c) 36 credits of courses in the Common Core Curriculum, comprising at least one and not more than two courses from each Area of Inquiry with not more than 24 credits of courses being selected within one academic year except where candidates are required to make up for failed credits; and
d) non-credit bearing courses as required by the University.

Capstone Experience
Students are required to complete the 6-credit "COMP3522 Real-life data science" and the 6-credit "COMP4501 Data science in discipline project" or "COMP4502 Final year project" to fulfill the capstone experience requirement for the degree of BEng in Data Science and Engineering.

Internship

Students are required to complete the non-credit bearing internship "COMP3510 Internship", which normally takes place after their third year of study.

The details of the distribution of the above course categories are as follows:

The curriculum of BEng(DS&E) comprises 240 credits of courses with the following structure:

UG 5 Requirements (54 credits)

Course Code	Course	No. of credits
CAES1000	Core University English	6
CAES9542	Technical English for Computer Science	6
CENG9001	Practical Chinese for Engineering Students	6
CC##XXXX	University Common Core Course (6 courses)*	36
XXXXxxxx	Non-credit bearing courses as required by the University	0
Total for UG5 Requirements		54

* Students have to complete 36 credits of courses in the Common Core Curriculum, comprising at least one and not more than two courses from each Area of Inquiry with not more than 24 credits of courses being selected within one academic year except where candidates are required to make up for failed credits.

Engineering Core Courses (24 credits)

Course Code	Course	No. of credits
ENGG1320	Engineers in the modern world	6
ENGG1330	Computer programming I	6
ENGG1340	Computer programming II	6
MATH1013	University mathematics II	6
Total for Engineering Core Courses		24

Discipline Core Courses (48 credits)

Introductory Courses (30 credits)

Course Code	Course	No. of credits
COMP2119	Introduction to data structure and algorithms	6
COMP2501	Introduction to data science and engineering	6
MATH2014	Multivariable calculus and linear algebra	6
STAT2601	Probability and statistics I	6
STAT2602	Probability and statistics II	6
Total for Introductory Discipline Core Courses		30

Advanced Courses (18 credits)

Course Code	Course	No. of credits
COMP3278	Introduction to database management systems	6
COMP3314	Machine learning	6
LLAWxxxx	Law and ethics in data science	6
Total for Advanced Discipline Core Courses		18

Capstone Experience and Internship (12 credits)

Course Code	Course	No. of credits
COMP3510	Internship*	0
COMP3522	Real-life data science[+]	6
COMP4501 or COMP4502	Data science in discipline project[+] or Final year project[+]	6
Total for Capstone Experience and Internship		12

* Students who are selected to participate in the Undergraduate Research Fellowship Programme are required to complete COMP3413 Research internship and are not required to complete COMP3510 Internship

[+]Capstone Experience

Discipline Elective Courses (30 credits)

at least 30 credits of courses to be chosen from the following list:

Course Code	Course	No. of credits
COMP3160	Web3.0 for Social Impact: An Innovative and Experiential Venture	6
COMP3270	Artificial intelligence	6
COMP3317	Computer vision	6
COMP3323 / FITE3010	Advanced database systems / Big data and data mining	6
COMP3340	Applied deep learning	6
COMP3353	Bioinformatics	6
COMP3355	Cyber security	6
COMP3361	Natural language processing	6
COMP3362	Hands-on AI: experimentation and applications	6
COMP3407	Scientific computing	6
COMP3513	Big data systems	6
COMP3516	Data analytics for IoT	6
COMP3520	Special topics in data science	6
COMP3521/ STAT3622	Visualization for data analytics/ Data Visualization	6
FITE2010	Distributed ledger and blockchain	6
SOWK3136	Application of big data analytics in social sciences	6
STAT3600	Linear statistical analysis	6
STAT3612	Statistical machine learning	6
STAT3621	Statistical data analysis	6
STAT4601	Time-series analysis	6
STAT4602	Multivariate data analysis	6

Elective Courses (72 credits)

At least 72 credits of courses offered by any department, except Common Core Courses.

Students are encouraged to pursue minor programme(s) related to application of data science. Recommended minor programmes: Finance, Economics, Marketing, Politics and Public Administration, Journalism and Media Studies, Social Data Science, Neuroscience, General Linguistics, Genetics and Genomics, Urban Studies, Urban Infrastructure Informatics, Industrial Engineering and Logistics Management, Earth Sciences, Environmental Science, Molecular Biology and Biotechnology.

Summary of curriculum structure of BEng in Data Science and Engineering

Course Categories	No. of credits
UG5 Requirements	54
Engineering Core Courses	24
Discipline Core Courses (Introductory)	30
Discipline Core Courses (Advanced)	18
Capstone Experience and Internship	12
Discipline Elective Courses	30
Elective Courses	72
Total	**240**

A sample study plan is given as follows:

FIRST YEAR

Engineering Core Courses (24 credits)		
ENGG1320	Engineers in the modern world	6
ENGG1330	Computer programming I	6
ENGG1340	Computer programming II	6
MATH1013	University mathematics II	6
Introductory Discipline Core Courses (12 credits)		
MATH2014	Multivariable calculus and linear algebra	6
COMP2501	Introduction to data science and engineering	6
University Requirements (UG5) (24 credits)		
CAES1000	Core University English	6
CC##XXXX	Three Common Core Courses	18

SECOND YEAR

Introductory Discipline Core Courses (18 credits)		
COMP2119	Introduction to data structure and algorithms	6
STAT2601	Probability and statistics I	6
STAT2602	Probability and statistics II	6
University Requirements (UG5) (18 credits)		
CC##XXXX	Three Common Core Course	18
Discipline Elective Courses (6 credits)		6
Elective Courses (18 credits)		18

THIRD YEAR

Advanced Discipline Core Courses (18 credits)		
COMP3278	Introduction to database management systems	6
COMP3314	Machine learning	6
LLAWxxxx	Law and ethics in data science	6

Capstone Experience (6 credits)		
COMP3522	Real-life data science	6
Internship (0 credit)		
COMP3510	Internship	0
University Requirements (UG5) (6 credits)		
CENG9001	Practical Chinese for Engineering Students (This course should be enrolled in the third year)	6
Discipline Elective Courses (12 credits)		12
Elective Courses (18 credits)		18

FOURTH YEAR

Discipline Elective Courses (12 credits)		12
Capstone Experience (6 credits)		
COMP4501 or COMP4502	Data science in discipline project or Final year project	6
University Requirements (UG5) (6 credits)		
CAES9542	Technical English for Computer Science	6
Elective Courses (36 credits)		36

Non-credit bearing courses as required by the University
Students will have the flexibility to take the courses in any semester throughout the period of studies.

COURSE DESCRIPTIONS

Candidates will be required to do the coursework in the respective courses selected. Not all courses are offered every semester.

Engineering Core Courses

ENGG1320. Engineers in the modern world (6 credits)

This course introduces fundamental concepts of engineering business; business models and financing; SWOT and market analysis; engineering entrepreneurship and innovation; system design, integration, and operation; product design and realization; and engineering sustainability. The course also involves hands-on projects in which students work in group to experience methods and techniques for the development of engineering business ideas and plans, products, or services.

Assessment: 100% continuous assessment

ENGG1330. Computer programming I (6 credits)

This is an introductory course designed for first-year engineering students to learn about computer

programming. Students will acquire basic Python programming skills, including syntax, identifiers, control statements, functions, recursions, strings, lists, dictionaries, tuples and filed. Searching and sorting algorithms, such as sequential search, binary search, bubble sort, insertion sort and selection sort, will also be covered.

Mutually exclusive with: COMP1117 or ENGG1111
Assessment: 70% continuous assessment, 30% examination

ENGG1340. Computer programming II (6 credits)

This course covers intermediate to advanced computer programming topics on various technologies and tools that are useful for software development. Topics include Linux shell commands, shell scripts, C/C++ programming, and separate compilation techniques and version control. This is a self-learning course; there will be no lecture and students will be provided with self-study materials. Students are required to complete milestone-based self-assessment tasks during the course. This course is designed for students who are interested in Computer Science /Computer Engineering.

Pre-requisite: ENGG1330 or COMP1117
Mutually exclusive with: COMP2113 or COMP2123
Assessment: 70% continuous assessment, 30% examination

MATH1013 University mathematics II (6 credits)

This course aims at students with Core Mathematics plus Module 1 or Core Mathematics plus Module 2 background and provides them with basic knowledge of calculus and some linear algebra that can be applied in various disciplines. It is expected to be followed by courses such as MATH2012, MATH2101, MATH2102, MATH2211, and MATH2241. Topics include: Functions; graphs; inverse functions; Limits; continuity and differentiability; Mean value theorem; Taylor's theorem; implicit differentiation; L'Hopital's rule; Higher order derivatives; maxima and minima; graph sketching; Radian, calculus of trigonometric functions; Definite and indefinite integrals; integration by substitutions; integration by parts; integration by partial fractions; Complex numbers, polar form, de Moivre's formula; Applications: Solving simple ordinary differential equations; Basic matrix and vector (of orders 2 and 3) operations, determinants of 2x2 or 3x3 matrices.

Prerequisite: Level 2 or above in Module 1, or Module 2 of HKDSE Mathematics or equivalent, or Pass in MATH1009 or MATH1011
Mutually exclusive with: MATH1821, or (MATH1851 and MATH1853)
Assessment: 50% continuous assessment, 50% examination

<u>University Requirements on Language Enhancement Courses</u>

CAES1000. Core University English (6 credits)
CENG9001. Practical Chinese for Engineering Students (6 credits)

Please refer to the University Language Enhancement Courses in the syllabus for the degree of BEng for details.

CAES9542. Technical English for Computer Science (6 credits)

Running alongside Computer Science, Financial Technology, Data Science related final-year / capstone

project courses, this one-semester, 6-credit course will build and consolidate students' ability to compose technical reports, and make technical oral presentations. The focus of this course is on helping students to report on the progress of their Final Year Project in an effective, professional manner in both written and oral communication. Topics include accessing, abstracting, analyzing, organizing and summarizing information; making effective grammatical and lexical choices; technical report writing; and technical presentations. Assessment is wholly by coursework.

Co-requisite: COMP3522 or COMP4801 or FITE4801 Assessment: 100% continuous assessment.

University Common Core Curriculum

Successful completion of 36 credits of courses in the Common Core Curriculum, comprising at least one and not more than two courses from each Area of Inquiry with not more than 24 credits of courses being selected within one academic year except where candidates are required to make up for failed credits:

- Science, Technology and Big Data
- Arts and Humanities
- Global Issues
- China: Culture, State and Society

Discipline Core Courses

COMP2119. Introduction to data structures and algorithms (6 credits)

Arrays, linked lists, trees and graphs; stacks and queues; symbol tables; priority queues, balanced trees; sorting algorithms; complexity analysis.

Prerequisite: COMP2113 or COMP2123 or ENGG1340
Assessment: 40% continuous assessment, 60% examination

COMP2501. Introduction to data science and engineering (6 credits)

The course introduces basic concepts and methodology of data science. The goal of this course is to provide students with an overview and practical experience of the entire data analysis process. Topics include: data source and data acquisition, data preparation and manipulation, exploratory data analysis, statistical and predictive analysis, data visualization and communication.

Prerequisite: COMP1117 or ENGG1330
Mutually exclusive with: STAT1005 or STAT1015
Assessment: 50% continuous assessment, 50% examination

MATH2014. Multivariable calculus and linear algebra (6 credits)

This course provides students with a solid foundation in calculus of several variables and linear algebra, which they will need in the study of mathematics related subjects. Topics include: Vectors and Matrices: Vectors in space, dot product and cross product, determinants (with geometric interpretations); Partial Derivatives: Functions of several variables, partial derivatives, extreme values and Lagrange multipliers, Taylor's formula; Multiple Integrals: Double and triple integrals, substitution in multiple integrals;

Matrix Algebra: Matrix addition and multiplication, system of linear equations as a matrix equation; Vector Spaces: The Euclidean spaces as vector spaces, its subspaces, span of vectors, linear independence, basis and dimension; Eigenvalues and Eigenvectors: Diagonalization and computing powers; Numerical Methods: Bisection method and Newton's method for finding roots of equations, Simpson's rule and Trapezoidal rule for numerical integration.

Prerequisite: MATH1013, or (MATH1851 and MATH1853)
Mutually exclusive with: MATH2822, or [(MATH2101 or MATH2102) and MATH2211]
Assessment: 5% assignments, 45% test, 50% examination

STAT2601. Probability and statistics I

The discipline of statistics is concerned with situations in which uncertainty and variability play an essential role and forms an important descriptive and analytical tool in many practical problems. Against a background of motivating problems this course develops relevant probability models for the description of such uncertainty and variability. Topics include: Sample spaces; Operations of events; Probability and probability laws; Conditional probability; Independence; Discrete random variables; Cumulative distribution function (cdf); Probability mass function (pmf); Bernoulli, binomial, geometric, and Poisson distributions; Continuous random variables; Cumulative distribution function (cdf); Probability density function (pdf); Exponential, Gamma, and normal distributions; Functions of a random variable; Joint distributions; Marginal distributions; Independent random variables; Functions of jointly distributed random variables; Expected value; Variance and standard deviation; Covariance and correlation.

Prerequisite/Co-requisite: MATH2014, or (MATH2101 and MATH2211)
Mutually exclusive with: ELEC2844 or MATH3603 or STAT1603 or STAT2901
Assessment: 40% continuous assessment, 60% examination

STAT2602. Probability and statistics II (6 credits)

This course builds on STAT2601, introducing further the concepts and methods of statistics. Emphasis is on the two major areas of statistical analysis: estimation and hypothesis testing. Through the disciplines of statistical modelling, inference and decision making, students will be equipped with both quantitative skills and qualitative perceptions essential for making rigorous statistical analysis of real-life data. Topics include: Overview: random sample; sampling distributions of statistics; moment generating function; large-sample theory: laws of large numbers and Central Limit Theorem; likelihood; sufficiency; factorisation criterion; Estimation: estimator; bias; mean squared error; standard error; consistency; Fisher information; Cramer-Rao Lower Bound; efficiency; method of moments; maximum likelihood estimator; Hypothesis testing: types of hypotheses; test statistics; p-value; size; power; likelihood ratio test; Neyman-Pearson Lemma; generalized likelihood ratio test; Pearson chi-squared test; Wald tests; Confidence interval: confidence level; confidence limits; equal-tailed interval; construction based on hypothesis tests.

Prerequisite: STAT2601
Mutually exclusive with: STAT3902
Assessment: 25% continuous assessment, 75% examination

COMP3278. Introduction to database management systems (6 credits)

This course studies the principles, design, administration, and implementation of database management systems. Topics include: entity-relationship model, relational model, relational algebra, database design and normalization, database query languages, indexing schemes, integrity and concurrency

control.

Prerequisite: COMP2119 or COMP2502 or ELEC2543 or FITE2000
Mutually exclusive with: IIMT3601
Assessment: 50% continuous assessment, 50% examination

COMP3314. Machine learning (6 credits)

This course introduces algorithms, tools, practices, and applications of machine learning. Topics include core methods such as supervised learning (classification and regression), unsupervised learning (clustering, principal component analysis), Bayesian estimation, neural networks; common practices in data pre-processing, hyper-parameter tuning, and model evaluation; tools/libraries/APIs such as scikit-learn, Theano/Keras, and multi/many-core CPU/GPU programming.

Prerequisites: MATH1853 or MATH2014; and COMP2119 or ELEC2543 or FITE2000
Assessment: 50% continuous assessment, 50% examination

LLAWxxxx. Law and ethics in data science (6 credits)

The primary objective of this course is to explore the legal and ethical challenges and ramifications in the modern practice of data science. Using a case-based approach, students will analyse contemporary controversies from a techno-legal and ethical perspectives. The focuses are data privacy and the regulation of using data in specific areas of law. Topics include basic privacy protection techniques, such as encryption and data anonymization data privacy laws, open data policy, data protection process and technology, issues in the usage of sensitive personal data and public data.

Assessment: 50% continuous assessment, 50% examination

<u>Discipline Elective Courses</u>

COMP3160. Web3.0 for Social Impact: An Innovative and Experiential Venture (6 credits)

This interdisciplinary course, jointly offered by the School of Computing and Data Science and the Department of Politics and Public Administration, explores how Web 3.0 technologies—including blockchain, AI, and decentralized systems—can drive meaningful social change through innovation and entrepreneurship. Students from computer science, AI, data science, and social science will collaborate to examine real-world applications of Web 3.0 in sustainability, economic development, community well-being, and governance. The first part of the course provides an accessible introduction to foundational Web 3.0 technologies, along with case studies from Hong Kong and beyond, while the second part focuses on practical implementation—covering product design, proof of concept development, financial and social valuation, fundraising, and project management. Through hands-on workshops led by Web 3.0 practitioners (developers, incubators, investors, and policymakers), students will gain firsthand insights and present their final projects to a panel of impact investors and domain experts.

Assessment: 100% continuous assessment

COMP3270. Artificial intelligence (6 credits)

This is an introduction course on the subject of artificial intelligence. Topics include: intelligent agents; search techniques for problem solving; knowledge representation; logical inference; reasoning under uncertainty; statistical models and machine learning.

Prerequisite: COMP2119 or FITE2000
Mutually exclusive with: ELEC4544 or IIMT3688
Assessment: 50% continuous assessment, 50% examination

COMP3317. Computer vision (6 credits)

This course introduces the principles, mathematical models and applications of computer vision. Topics include: image processing techniques, feature extraction techniques, imaging models and camera calibration techniques, stereo vision, and motion analysis.

Prerequisites: COMP2119; and MATH1853 or MATH2014 or MATH2101
Assessment: 50% continuous assessment, 50% examination

COMP3323. Advanced database systems (6 credits)

The course will study some advanced topics and techniques in database systems, with a focus on the system and algorithmic aspects. It will also survey the recent development and progress in selected areas. Topics include: query optimization, spatial-spatiotemporal data management, multimedia and time-series data management, information retrieval and XML, data mining.

Prerequisite: COMP3278
Mutually exclusive with: FITE3010
Assessment: 50% continuous assessment, 50% examination

COMP3340. Applied deep learning (6 credits)

An introduction to algorithms and applications of deep learning. The course helps students get hands-on experience of building deep learning models to solve practical tasks including image recognition, image generation, reinforcement learning, and language translation. Topics include: machine learning theory; optimization in deep learning; convolutional neural networks; recurrent neural networks; generative adversarial networks; reinforcement learning; self-driving vehicle.

Prerequisites: COMP2119 or ELEC2543 or FITE2000; and MATH1853 or MATH2014
Mutually exclusive with: ELEC4544
Assessment: 50% continuous assessment, 50% examination

COMP3353. Bioinformatics (6 credits)

The goal of the course is for students to be grounded in basic bioinformatics concepts, algorithms, tools, and databases. Students will be leaving the course with hands-on bioinformatics analysis experience and empowered to conduct independent bioinformatics analyses. We will study: 1) algorithms, especially those for sequence alignment and assembly, which comprise the foundation of the rapid development of bioinformatics and DNA sequencing; 2) the leading bioinformatics tools for comparing and analyzing genomes starting from raw sequencing data; 3) the functions and organization of a few essential bioinformatics databases and learn how they support various types of bioinformatics analysis.

Prerequisite: COMP1117 or ENGG1330
Assessment: 70% continuous assessment, 30% examination

COMP3355. Cyber security (6 credits)

This course introduces the principles, mechanisms and implementation of cyber security and data protection. Knowledge about the attack and defense are included. Topics include notion and terms of cyber security; network and Internet security, introduction to encryption: classic and modern encryption technologies; authentication methods; access control methods; cyber attacks and defenses (e.g. malware, DDoS).

Prerequisite: COMP2119 or ELEC2543 or FITE2000
Mutually exclusive with: ELEC4641
Assessment: 50% continuous assessment, 50% examination

COMP3361. Natural language processing (6 credits)

Natural language processing (NLP) is the study of human language from a computational perspective. The course will be focusing on machine learning and corpus-based methods and algorithms. We will cover syntactic, semantic and discourse processing models. We will describe the use of these methods and models in applications including syntactic parsing, information extraction, statistical machine translation, dialogue systems, and summarization. This course starts with language models (LMs), which are both front and center in natural language processing (NLP), and then introduces key machine learning (ML) ideas that students should grasp (e.g. feature-based models, log-linear models and then the neural models). We will land on modern generic meaning representation methods (e.g. BERT/GPT-3) and the idea of pretraining / finetuning.

Prerequisites: COMP3314 or COMP3340; and MATH1853
Assessment: 50% continuous assessment, 50% examination

COMP3362. Hands-on AI: experimentation and applications (6 credits)

This course comprises two main components: students first acquire the basic know-how of the state-of-the-art AI technologies, platforms and tools (e.g., TensorFlow, PyTorch, scikit-learn) via example-based modules in a self-paced learning mode. Students will then identify a creative or practical data-driven application and implement an AI-powered solution for the application as the course project. Students will be able to experience a complete AI experimentation and evaluation cycle throughout the project.

Prerequisite: COMP3314
Mutually exclusive with: COMP3359
Assessment: 100% continuous assessment

COMP3407. Scientific computing (6 credits)

This course provides an overview and covers the fundamentals of scientific and numerical computing. It focuses topics in numerical analysis and computation, with discussions on applications of scientific computing.

Prerequisites: COMP1117 or ENGG1330; and COMP2121
Assessment: 50% continuous assessment, 50% examination

COMP3513. Big data systems (6 credits)

The objective of this course is to study the design and implementation of Big Data systems. Topics include: data analytics pipelines, data processing framework, distributed and parallel data systems, network attached storage, data storage virtualization, query language support, data center architecture, fault tolerance, and recovery.

Prerequisites: COMP2501; and COMP3278
Assessment: 50% continuous assessment, 50% examination

COMP3516. Data analytics for IoT (6 credits)

This course introduces basic concepts, technologies, and applications of the Internet of Things (IoT), with a focus on data analytics. The course covers a range of enabling techniques in sensing, computing, analytics, learning for IoT and connects them to exciting applications in smart homes, healthcare, security, etc. The lectures cover the pipeline of data generation, data acquisition, data transportation, data analysis and learning, and data applications, with various topics from the fundamentals (e.g., signal processing, statistical analysis, machine learning) to real-world systems. Billions of things are connected today, and this course helps students to understand how IoT will evolve into AIoT (Artificial Intelligence of Things).

Prerequisite: COMP2119
Assessment: 60% continuous assessment, 40% examination

COMP3520. Special topics in data science (6 credits)

Data science is an emerging area. The primary objective of this course is to introduce new development in this area, including but not limited to advanced computational techniques, latest advances in technologies related to data science, and challenging R&D problems. Selected topics in data science that are of current interest will be discussed. Topics may vary from year to year.

Prerequisites: MATH1013; and COMP1117; and COMP2113
Assessment: 50% continuous assessment, 50% examination

COMP3521. Visualization for data analytics (6 credits)

This course aims to give an overview of the basic principles and techniques for visualization and visual analytics. In particular, topics including human visual perception, color and visualization techniques for various data kinds (e.g., spatial, geospatial and multivariate data, graphs and networks, text and document) will be covered. The use of interactive visual interface to facilitate analytical reasoning will also be discussed. Students will use practical tools and apply visualization principles and techniques to perform visual data analysis on large datasets.

Prerequisite: COMP2119 or COMP2502 or ELEC2543 or FITE2000
Mutually exclusive with: STAT3622

Assessment: 50% continuous assessment, 50% examination

FITE2010. Distributed ledger and blockchain (6 credits)

This course introduces basic theories of blockchain and distributed ledger, which includes basic cryptography, public key cryptosystem, distributed computing and consensus protocols. Financial applications of blockchain and distributed ledger will be discussed.

Prerequisites: FITE1010 or MATH1853 or MATH2101; and COMP2119 or ELEC2543 or FITE2000
Assessment: 40% continuous assessment, 60% examination

FITE3010. Big data and data mining (6 credits)

The goal of the course is to study the main methods used today for data mining and on-line analytical processing. Topics include Big Data Architecture, Data Mining Algorithms, Classification, and Clustering.

Prerequisites: FITE1010 or MATH1853 or MATH2101; and COMP2119 or ELEC2543 or FITE2000
Mutually exclusive with: COMP3323
Assessment: 40% continuous assessment, 60% examination

SOWK3136. Application of big data analytics in social sciences (6 credits)

Do Google and Facebook understand us better than we know ourselves? Are we being descended to lab rats every time we go online? Can we extract information from electronic health records to prevent diseases or even suicide? Is the impartially designed algorithm for predicting an individual's probability of recidivism truly fair for sentencing individuals who have committed crimes? When big data analytics are routinely applied to nudging our daily lives, the ability to audit the algorithms adopted by these analytics becomes crucial.

The course will focus on elaborating the core principles of a variety of techniques adopted when predicting future phenomena through the lens of big data. We will use a case study approach to provide an in-depth understanding of how predictions are made using various big data analytics. Students will be guided to develop a rich contextual understanding of consequences associated with applications of big data in different scenarios. The goal of this course is to inspire the students to think creatively and critically about how big data analytics can be used to making scientific discoveries and doing social good. Meanwhile, they will also learn to identify potential prejudices embedded in poorly designed algorithms and be able to stand up against the abuse of big data.

Assessment: 100% coursework.

STAT3600. Linear statistical analysis (6 credits)

The analysis of variability is mainly concerned with locating the sources of the variability. Many statistical techniques investigate these sources through the use of 'linear' models. This course presents the theory and practice of these models. Topics include: Simple linear regression: least squares method, analysis of variance, coefficient of determination, hypothesis tests and confidence intervals for regression parameters, prediction; Multiple linear regression: least squares method, analysis of variance,

coefficient of determination, reduced vs full models, hypothesis tests and confidence intervals for regression parameters, prediction, polynomial regression; One-way classification models: one-way ANOVA, analysis of treatment effects, contrasts; Two-way classification models: interactions, two-way ANOVA for balanced data structures, analysis of treatment effects, contrasts, randomised complete block design; Universal approach to linear modelling: dummy variables, 'multiple linear regression' representation of one-way and two-way (unbalanced) models, ANCOVA models, concomitant variables; Regression diagnostics: leverage, residual plot, normal probability plot, outlier, studentized residual, influential observation, Cook's distance, multicollinearity, model transformation.

Prerequisite: STAT2602
Mutually exclusive with: STAT3907
Assessment: 25% continuous assessment, 75% examination

STAT3612. Statistical machine learning (6 credits)

Machine learning is the study of computer algorithms that build models of observed data in order to make predictions or decisions. Statistical machine learning emphasizes the importance of statistical theory and methodology in the algorithmic development. This course provides a comprehensive and practical coverage of essential machine learning concepts and a variety of learning algorithms under supervised and unsupervised settings. The course materials are presented with lots of examples and reproducible codes. Topics include: Data science, data exploration, generalized linear models, variable selection, basis expansion, regularization, cross-validation, tree-based methods, kernel methods, neural networks, dimension reduction, principal component analysis, cluster analysis, stochastic optimization, interpretable machine learning.

Prerequisites: STAT2602, or (STAT1603 and any University level 2 course) or STAT3902; and STAT3600 or STAT3907
Mutually exclusive with: STAT4904
Assessment: 100% continuous assessment

STAT3621. Statistical data analysis (6 credits)

Building on prior coursework in statistical methods and modeling, students will get a deeper understanding of the entire process of data analysis. The course aims to develop skills of model selection and hypotheses formulation so that questions of interest can be properly formulated and answered. An important element deals with model review and improvement, when one's first attempt does not adequately fit the data. Students will learn how to explore the data, to build reliable models, and to communicate the results of data analysis to a variety of audiences. Topics include: Descriptive statistics, presentation and visualization of data; Simple statistical analyses for the one-sample and two-sample case using parametric and nonparametric methods; Regression analyses: model fitting; variable selection and model diagnostic checking; Analysis of Variance (ANOVA): 1-way, two-way and higher-way ANOVA; Covariance analysis; Categorical and count data: binary logistic regression, Poisson regression. Real data sets will be presented for modelling and analysis using statistical software for gaining hands-on experience.

Prerequisite: STAT3600 or STAT3907
Assessment: 50% continuous assessment, 50% examination

STAT4601. Time-series analysis (6 credits)

A time series consists of a set of observations on a random variable taken over time. Time series arise naturally in climatology, economics, environment studies, finance and many other disciplines. The observations in a time series are usually correlated; the course establishes a framework to discuss this. This course distinguishes different type of time series, investigates various representations for the processes and studies the relative merits of different forecasting procedures. Students will analyse real time-series data on the computer. Topics include: Stationarity and the autocorrelation functions; linear stationary models; linear non-stationary modes; model identification; estimation and diagnostic checking; seasonal models and forecasting methods for time series.

Prerequisite: STAT3600
Mutually exclusive with: STAT3614, STAT3907
Assessment: 40% continuous assessment, 60% examination

STAT4602. Multivariate data analysis (6 credits)

In many designed experiments or observational studies, the researchers are dealing with multivariate data, where each observation is a set of measurements taken on the same individual. These measurements are often correlated. The correlation prevents the use of univariate statistics to draw inferences. This course develops the statistical methods for analysing multivariate data through examples in various fields of application and hands-on experience with the statistical software SAS. Topics include: Problems with multivariate data. Multivariate normality and transforms. Mean structure for one sample. Tests of covariance matrix. Correlations: Simple, partial, multiple and canonical. Multivariate regression. Principal components analysis. Factor analysis. Problems for means of several samples. Multivariate analysis of variance. Discriminant analysis. Classification. Multivariate linear model.

Prerequisite: STAT3600 or STAT3907
Assessment: 40% continuous assessment, 60% examination

Capstone Experience and Internship

COMP3510. Internship (0 credit)

The course consists of two components: internship and professionalism. Internship requires students to spend a minimum of four weeks employed, full-time, as IT interns or trainees. During this period, they are engaged in work of direct relevance to their programme of study. The Internship provides students with practical, real-world experience and represents a valuable complement to their academic training. Professionalism exposes students to social and professional issues in computing. Students need to understand their professional roles when working as data science professionals as well as the responsibility that they will bear. They also need to develop the ability to ask serious questions about the social impact of data science and engineering and to evaluate proposed answers to those questions. Topics include: intellectual property, privacy, social context of computing, risks, safety and security concerns for data science professionals, professional and ethical responsibilities, and continuing professional development.

Assessment: 100% continuous assessment

COMP3522. Real-life data science (6 credits)

In this course, students will learn data science step by step through real analytics example: data mining, modelling, tableau visualization and more. Unlike many classes where everything works just the way it should and the training is smooth sailing, this course will give students a data science odyssey through experiencing the pains a data scientist goes through on a daily basis. Corrupt data, anomalies, irregularities, etc. Upon completing this course, the students will enhance their data wrangling skills and learn how to 1) model their data, 2) curve-fit their data, and 3) how to communicate their findings. The students will develop a good understanding of Tableau, SQL, SSIS, and Gretl that give them a safe ride in data lakes. With no final exam, the students will be given practical exercises that prepare them to be at the helm for real-world challenges.

Prerequisite: ENGG1330
Assessment: 100% continuous assessment

COMP4501. Data science in discipline project (6 credits)

Students will work in groups or individually on a capstone project which is on data science in association with a domain focus. Students are required to identify a data-intensive problem in a specific application domain, and to implement a data-driven solution for the problem. Students will undergo a complete data science project life cycle, from problem understanding, data collection, data exploration to data modelling, analysis and interpretation, and finally deliver a data science solution.

Mutually exclusive with: COMP4502
Assessment: 100% continuous assessment

COMP4502. Final year project (6 credits)

Student individuals or groups, during the final year of their studies, undertake full end-to-end development of a substantial project, taking it from initial concept through to final delivery. Topics range from applied technologies to assignments on basic research in relation to data science and engineering. In case of a team project, significant contribution is required from each member and students are assessed individually. Strict standards of quality will be enforced throughout the project development.

Mutually exclusive with: COMP4501
Assessment: 100% continuous assessment
