Master of
Data Science

Apply now for entry in September 2019

Machine Learning

Data Analytics

Advanced Statistical Modelling

Social Networks

Cloud Computing

THE UNIVERSITY OF HONG KONG
FACULTY OF SCIENCE

The amount and complexity of digital data have grown exponentially over the years. In answer to the desperate call for experts in processing such data, Department of Statistics and Actuarial Science, The University of Hong Kong, has recently launched a new Taught Master Programme, Master of Data Science (MDASC), jointly with the Department of Computer Science, to teach students how to analyze data and formulate data-driven strategies.

Programme Information

Master of Data Science (MDASC) is a taught master programme jointly offered by Department of Statistics and Actuarial Science (host) and Department of Computer Science.

Its interdisciplinarity promotes the applications of computer technology, operational research, statistical modelling, and simulation to decision-making and problem-solving in all organizations and enterprises within the private and public sectors.

The curriculum of the MDASC programme adopts a well-balanced and comprehensive pedagogy of both statistical and computational concepts and methodologies, underpinning applications that are not limited to business or a single field alone.

It is a programme ideal for

1. those whose interest in high-level analytical skills straddles the disciplinary divide between statistics and computational analytics, and

2. those who wish to pursue further study in the field of data science after studying science, social sciences, engineering, medical sciences, information systems, computing and data analytics in their undergraduate studies.

Reimbursable Course(s) by Continue Education Fund (CEF)

Certain course(s) of the programme would be included in the list of reimbursable course for CEF purposes.

持續進修基金
Continuing Education Fund

Programme Highlights

- Joint programme offered by Department of Statistics and Actuarial Science and Department of Computer Science
- Interdisciplinary and comprehensive curriculum
- Solid foundation in statistical and computational analyses
- Students can select electives from Computer Science, Mathematics and Statistics
- Electives cover a broad range of contemporary topics
- Hands-on applications of methodologies with powerful software
- Capstone project with real-life scenario

Course Highlights

The core courses of the proposed MDASC programme mainly focus on both predictive and prescriptive concepts and methodologies with an effort to equip students with a solid foundation in statistical and computational analyses, e.g.

Data Science technology   Cluster & cloud computing
Data mining techniques   Big data analytics

The electives cover a broad range of contemporary topics and provide students with solid training in diverse and applied techniques used in data science, including but not limited to

Financial data analysis   Marketing analytics   Quantitative risk
Management and finance   Network security
Computational intelligence   Machine learning
Multimedia technologies   Smart phone apps development

"In the 21st century, we are increasingly reliant on technology for day-to-day living. The best jobs of 2018 reflect this ongoing shift, with careers in tech or directly tied to tech represented. Mathematician and Data Scientist are two examples of the latter."

Extract from the Best Jobs of 2018 from CareerCast.com

"Integration of statistical inference principles as part of Big Data will be essential to resolve these (big data) challenges."

Extracted from the Federal Big Data Research and Development Strategic Plan by the Executive Office of the President of the USA.

Programme Curriculum

Commencing in September, the curriculum is composed of 72 credits of courses. Courses with 6 credits are offered in the first and second semesters while courses with 3 credits are normally offered in the summer semester. If a student selects a course whose contents are similar to a course (or courses) which he/she has taken in his/her previous study, the Department may not approve the selection in question. The curriculum is the same for both full-time and part-time study mode.

Compulsory Courses (36 credits)

COMP7305	Cluster and Cloud Computing (6 credits)	
COMP7404	Computational Intelligence and Machine Learning (6 credits)	
DASC7011	Statistical Inference for Data Science (6 credits)	
DASC7104	Advanced Database Systems (6 credits)	
STAT6014	Advanced Statistical Modelling (6 credits)	
STAT7008	Programming for Data Science (6 credits)	

Disciplinary Electives (24 credits)*

with at least 12 credits from List A and 12 credits from List B

List A

COMP7105	Advanced Topics in Data Science (6 credits) **NEW**
COMP7503	Multimedia Technologies (6 credits)
COMP7506	Smart Phone Apps Development (6 credits)
COMP7507	Visualization and Visual Analytics (6 credits)
COMP7906	Introduction to Cyber Security (6 credits)
DASC7606	Deep Learning (6 credits)
ICOM6044	Data Science for Business (6 credits)

List B

MATH8502	Topics in Applied Discrete Mathematics (6 credits) **NEW**
MATH8503	Topics in Mathematical Programming and Optimization (6 credits) **NEW**
STAT6013	Financial Data Analysis (6 credits)
STAT6015	Advanced Quantitative Risk Management and Finance (6 credits)
STAT6016	Spatial Data Analysis (6 credits)
STAT8003	Time Series Forecasting (6 credits)
STAT8017	Data Mining Techniques (6 credits)
STAT8019	Marketing Analytics (6 credits)
STAT8021	Big Data Analytics (6 credits) **NEW**
STAT8300	Career Development and Communication Workshop (3 credits)
STAT8306	Statistical Methods for Network Data (3 credits)
STAT8307	Text Analytics (3 credits) **NEW**

Students who have completed the same courses in their previous studies in HKU, e.g. Master of Statistics or Master of Science in Computer Science may, on production of relevant transcripts, be permitted to select up to 24 credits of disciplinary electives from either List A or List B above if they are not able to find any untaken options from either of the lists of disciplinary electives.

Capstone requirement (12 credits)

DASC7600	Data Science Project (12 credits)	

Course Description

Compulsory Courses

COMP7305   Cluster and cloud computing (6 credits)

This course offers an overview of current cluster and cloud technologies, and discusses various issues in the design and implementation of cluster and cloud systems. Topics include cluster architecture, cluster middleware, and virtualization techniques (e.g., Xen, KVM) used in modern data centers. We will discuss three types of Cloud computing platforms, including SaaS, PaaS, and IaaS, by providing motivating examples from companies such as Google, Amazon, and Microsoft; and introduce Hadoop MapReduce and Spark programming paradigms for large-scale data analysis.

Prerequisites: The students are expected to exercise the systems configuration and administration under a Linux cluster. Basic understanding of Linux operating system and some experiences in system level programming (C/C++ or Java) are required.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

COMP7404   Computational intelligence and machine learning (6 credits)

This course will teach a broad set of principles and tools that will provide the mathematical and algorithmic framework for tackling problems using Artificial Intelligence (AI) and Machine Learning (ML). AI and ML are highly interdisciplinary fields with impact in different applications, such as, biology, robotics, language, economics, and computer science. AI is the science and engineering of making intelligent machines, especially intelligent computer programs, while ML refers to the changes in systems that perform tasks associated with AI.

Topics may include a subset of the following: problem solving by search, heuristic (informed) search, constraint satisfaction, games, knowledge-based agents, supervised learning, unsupervised learning; learning theory, reinforcement learning and adaptive control.

Pre-requisites: Nil, but knowledge of data structures and algorithms, probability, linear algebra, and programming would be an advantage.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

DASC7011   Statistical inference for data science (6 credits)

Computing power has revolutionized the theory and practice of statistical inference. Reciprocally, novel statistical inference procedures are becoming an integral part of data science. By focusing on the interplay between statistical inference and methodologies for data science, this course reviews the main concepts underpinning classical statistical inference, studies computer-intensive methods for conducting statistical inference, and examines important issues concerning statistical inference drawn upon modern learning technologies. Contents include classical frequentist and Bayesian inferences, computer-intensive methods such as the EM algorithm, the bootstrap and the Markov chain Monte Carlo, large-scale hypothesis testing, high-dimensional modeling, and post-model-selection inference.

Assessment: One 2-hour written examination; 40% coursework and 60% examination

DASC7104   Advanced database systems (6 credits)

The course will study some advanced topics and techniques in database systems, with a focus on the aspects of big data analytics, algorithms, and system design & organisation. 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.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

STAT6014   Advanced statistical modelling (6 credits)

This course introduces modern methods for constructing and evaluating statistical models and their implementation using popular computing software, such as R or Python. It will cover both the underlying principles of each modelling approach and the model estimation procedures. Topics from: (i) Generalized linear models; (ii) Mixed models; (iii) Kernel and local polynomial regression; (iv) Generalized additive models; (v) Hidden Markov models and Bayesian networks.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

STAT7008   Programming for data science (6 credits)

In the big data era, it is very easy to collect huge amounts of data. Capturing and exploiting the important information contained within such datasets poses a number of statistical challenges. This course aims to provide students with a strong foundation in computing skills necessary to use R or Python to tackle some of these challenges. Possible topics to be covered may include exploratory data analysis and visualization, collecting data from a variety of sources (e.g. excel, web-scraping, APIs and others), object-oriented programming concepts and scientific computation tools. Students will learn to create their own R packages or Python libraries.

Assessment: 100% coursework

Disciplinary Electives

COMP7105   Advanced topics in data science (6 credits)   NEW

This course will introduce selected advanced computational methods and apply them to problems in data analysis and relevant applications.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

COMP7503   Multimedia technologies (6 credits)

This course presents fundamental concepts and emerging technologies for multimedia computing. Students are expected to learn how to develop various kinds of media communication, presentation, and manipulation techniques. At the end of course, students should acquire proper skill set to utilize, integrate and synchronize different information and data from media sources for building specific multimedia applications. Topics include media data acquisition methods and techniques; nature of perceptually encoded information; processing and manipulation of media data; multimedia content organization and analysis; trending technologies for future multimedia computing.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

COMP7506   Smart phone apps development (6 credits)

Smart phones have become very popular in recent years. For iPhones alone, CEO Tim Cook announced that Apple has sold the billionth iPhone in July 2016. In addition to iPhones, there are also Android phones, Symbian phones as well as Windows phones.

Smart phones play an important role in mobile communication and applications. Smart phones are powerful as they support a wide range of applications (called apps). Most of the time, smart phone users just purchase their favorite apps wirelessly from the vendors. There is a great potential for software developer to reach worldwide users.

This course aims at introducing the design issues of smart phone apps. For examples, the smart phone screen is usually much smaller than the computer monitor. We have to pay special attention to this aspect in order to develop attractive and successful apps. Different smart phone apps development environments and programming techniques (such as Java for Android phones, Objective-C and Swift for iPhones) will be introduced to facilitate students to develop their own apps.

Prerequisites: Students should have basic programming knowledge, e.g. C++ or Java.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

COMP7507   Visualization and visual analytics (6 credits)

This course introduces the basic principles and techniques in visualization and visual analytics, and their applications. Topics include human visual perception; color; visualization techniques for spatial, geospatial and multivariate data, graphs and networks; text and document visualization; scientific visualization; interaction and visual analysis.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

COMP7906   Introduction to cyber security (6 credits)

The aim of the course is to introduce different methods of protecting information and data in the cyber world, including the privacy issue. Topics include introduction to security; cyber attacks and threats; cryptographic algorithms and applications; network security and infrastructure.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

DASC7606 Deep learning (6 credits)

Machine learning is a fast growing field in computer science and deep learning is the cutting edge technology that enables machines to learn from large-scale and complex datasets. This course will first introduce fundamental machine learning techniques and will then focus on artificial neural networks and how to train and optimize them to solve challenging problems using deep learning. Topics covered include linear and logistic regression, neural networks, convolutional neural networks, deep reinforcement learning and unsupervised feature learning. Popular deep learning software, such as Caffe, Torch and TensorFlow, will also be introduced.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

ICOM6044 Data science for business (6 credits)

The emerging discipline of data science combines statistical methods with computer science to solve problems in applied areas. In this case we focus on how data science can be used to solve business problems especially those in electronic commerce. By its very nature e-commerce is able to generate large amounts of data and data mining methods are quite helpful for managers in turning this data into knowledge which in turn can be used to make better decisions. These data sets and their accompanying quantitative methods have the potential to dramatically change decision making in many areas of business. For example, ideas like interactive marketing, customer relationship management, and database marketing are pushing companies to utilize the information they collect about their customers in order to make better marketing decisions.

This course focuses on how data science methods can be applied to solve managerial problems in marketing and electronic commerce. Our emphasis is developing a core set of principles that embody data science: empirical reasoning, exploratory and visual analysis, and predictive modeling. We use these core principles to understand many methods used in data mining and machine learning. Our strategy in this course is to survey several popular techniques and understand how they map into these core principles. These techniques are illustrated with case studies. However, the emphasis is not on the software for implementing these techniques but on understanding the inputs and outputs of these techniques and how they are used to solve business problems.

Pre-requisites: Students should not be taking or have taken STAT8017 Data mining techniques or equivalent

Assessment: One 2-hour written examination; 65% coursework and 35% examination

MATH8502 Topics in applied discrete mathematics (6 credits)
NEW

This course aims to provide students with the opportunity to study some further topics in applied discrete mathematics. A selection of topics in discrete mathematics applied in combinatorics and optimization (such as algebraic coding theory, cryptography, discrete optimization, etc.) The selected topics may vary from year to year.

Pre-requisites: Knowledge in introductory discrete mathematics. Students may be asked to present appropriate evidence of having met the pre-requisites for enrolling in this course.

Assessment: One 2.5-hour written examination; 50% coursework and 50% examination

MATH8503 Topics in mathematical programming and optimization (6 credits) NEW

A study in greater depth of some special topics in mathematical programming or optimization. It is mainly intended for students in Operations Research or related subject areas. This course covers a selection of topics which may include convex, quadratic, geometric, stochastic programming, or discrete combinatorial optimization. The selected topics may vary from year to year.

Pre-requisites: Knowledge in introductory mathematical programming and optimization. Students may be asked to present appropriate evidence of having met the pre-requisites for enrolling in this course.

Assessment: One 2.5-hour written examination; 50% coursework and 50% examination

STAT6013 Financial data analysis (6 credits)

This course aims at introducing statistical methodologies in analyzing financial data. Financial applications and statistical methodologies are intertwined in all lectures. Contents include: recent advances in modern portfolio theory, Copula, market microstructure and high frequency data analysis.

Assessment: One 2-hour written examination; 40% coursework and 60% examination

STAT6015 Advanced quantitative risk management and finance (6 credits)

This course covers statistical methods and models of importance to risk management and finance and links finance theory to market practice via statistical modelling and decision making. Emphases will be put on empirical analyses to address the discrepancy between finance theory and market data. Contents include: Elementary Stochastic Calculus; Basic Monte Carlo and Quasi-Monte Carlo Methods; Variance Reduction Techniques; Simulating the value of options and the value-at-risk for risk management; Review of univariate volatility models; multivariate volatility models; Value-at-risk and expected shortfall; estimation, back-testing and stress testing; Extreme value theory for risk management.

Assessment: One 2-hour written examination; 25% coursework and 75% examination

STAT6016 Spatial data analysis (6 credits)

This course covers statistical concepts and tools involved in modelling data which are correlated in space. Applications can be found in many fields including epidemiology and public health, environmental sciences and ecology, economics and others. Covered topics include: (1) Outline of three types of spatial data: point-level (geostatistical), areal (lattice), and spatial point process. (2) Model-based geostatistics: covariance functions and the variogram; spatial trends and directional effects; intrinsic models; estimation by curve fitting or by maximum likelihood; spatial prediction by least squares, by simple and ordinary kriging, by trans-Gaussian kriging. (3) Areal data models: introduction to Markov random fields; conditional, intrinsic, and simultaneous autoregressive (CAR, IAR, and SAR) models. (4) Hierarchical modelling for univariate spatial response data, including Bayesian kriging and lattice modelling. (5) Introduction to simple spatial point processes and spatio-temporal models. Real data analysis examples will be provided with dedicated R packages such as geoR.

Assessment: One 2-hour written examination; 50% coursework and 50% examination

STAT8003　Time series forecasting (6 credits)

A time series consists of a set of observations on a random variable taken over time. Such series arise naturally in climatology, economics, finance, environmental research and many other disciplines. In additional to statistical modelling, the course deals with the prediction of future behaviour of these time series. This course distinguishes different types of time series, investigates various representations for them and studies the relative merits of different forecasting procedures.

Assessment: One 3-hour written examination; 40% coursework and 60% examination

STAT8017　Data mining techniques (6 credits)

With the rapid developments in computer and data storage technologies, the fundamental paradigms of classical data analysis are mature for change. Data mining techniques aim at helping people to work smarter by revealing underlying structure and relationships in large amounts of data. This course takes a practical approach to introduce the new generation of data mining techniques and show how to use them to make better decisions. Topics include data preparation, feature selection, association rules, decision trees, bagging, random forests and gradient boosting, cluster analysis, neural networks, introduction to text mining.

Pre-requisites: Students should not be taking or have taken ICOM6044 Data science for business or equivalent

Assessment: 100% coursework

STAT8019　Marketing analytics (6 credits)

This course aims to introduce various statistical models and methodology used in marketing research. Special emphasis will be put on marketing analytics and statistical techniques for marketing decision making including market segmentation, market response models, consumer preference analysis and conjoint analysis. Contents include market response models, statistical methods for segmentation, targeting and positioning, statistical methods for new product design.

Assessment: One 3-hour written examination; 40% coursework and 60% examination

STAT8021　Big data analytics (6 credits) NEW

The recent explosion of social media and the computerization of every aspect of life resulted in the creation of volumes of mostly unstructured data (big data): web logs, e-mails, video, speech recordings, photographs, tweets and others. This course aims to provide students with knowledge and skills of some advanced analytics and statistical modelling for solving big data problems. Students are required to possess basic understanding of R language.

Pre-requisites: Pass in STAT8017 Data mining techniques or equivalent; students should not be taking or have taken STAT8307 Text analytics or equivalent

Assessment: 100% coursework

STAT8300　Career development and communication workshop (3 credits)

The course is specially designed for students who wish to sharpen their communication and career preparation skills through a variety of activities including lectures, skill-based workshops, small group discussion and role plays. All of which aim to facilitate students in making informed career choices, provide practical training to enrich communication, presentation, time management and advanced interview skills, and to enhance students' overall competitiveness in the employment markets.

Assessment: 100% coursework, assessment of this course is on a pass or fail or distinction basis

STAT8306　Statistical methods for network data (3 credits)

The six degree of separation theorizes that human interactions could be easily represented in the form of a network. Examples of networks include router networks, the World Wide Web, social networks (e.g. Facebook or Twitter), genetic interaction networks and various collaboration networks (e.g. movie actor coloration network and scientific paper collaboration network). Despite the diversity in the nature of sources, the networks exhibit some common properties. For example, both the spread of disease in a population and the spread of rumors in a social network are in sub-logarithmic time. This course aims at discussing the common properties of real networks and the recent development of statistical network models. Topics may include common network measures, community detection in graphs, preferential attachment random network models, exponential random graph models, models based on random point processes and the hidden network discovery on a set of dependent random variables.

Assessment: One 1.5-hour written examination; 50% coursework and 50% examination

STAT8307　Text analytics (3 credits) NEW

The textual data constitutes and enormous proportion of unstructured data which is characterized as one of 'V's in Big Data. The logical and computational reasonings are applied to transform large collection of written resources to structured data for use in further analysis, visualization, integration with structured data in database or warehouse, and further refinement using machine learning systems. This course introduces the methodology of text analytics. Topics are selected from natural language processing, word representation, text categorization and clustering, topic modelling and sentiment analysis. Students are required to possess basic understanding of R language.

Pre-requisites: Pass in STAT8017 Data mining techniques or equivalent; students should not be taking or have taken STAT8021 Big data analytics or equivalent

Assessment: 100% coursework

Capstone Requirement

DASC7600　Data science project (12 credits)

Candidate will be required to carry out independent work on a major project under the supervision of individual staff member. A written report is required.

Assessment: 75% written report and 25% oral presentation

Programme Duration

The programme normally extends over 1.5 academic years for full-time study, and 2.5 academic years for part-time study. Fast track completion is possible for both study modes, and students could graduate in 1 year for full-time study mode, and 2 years for part-time study mode. All lectures are conducted in English at HKU.

Optional Summer Courses

Tutorials on statistical software (e.g. R and SAS) and computer programming (e.g. Python) will be held in August, 2019 for students who need to rejuvenate their skills.
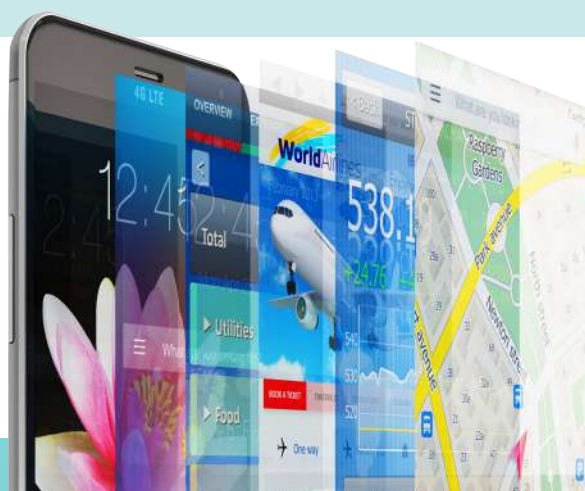
Data Science helps us decipher the data and extract valuable information that can be used as a strategic part of critical decision-making. It is therefore no surprise that Data Scientists are of huge demand by employers all over the world.

As an international premier centre par excellence, HKU always strives to provide students with a state-of-the-art education. The launch of Master of Data Science (MDASC) provides students with the chance to acquire the quantitative and analytical skills much needed to stay ahead of the game.

Testimonials

As a quantitative researcher, I have to apply a lot of knowledge in finance, risk management and statistics in my job. Therefore, after pursuing the FRM and CFA qualification, I started my MStat study in 2012 in order to strengthen my statistical knowledge in the area of finance and risk management. I learned a lot of useful statistical techniques during my MStat study.

Since practice is the sole criterion for verifying the truth, the best way to practice the theories and techniques I learned from MStat is to implement them in my job. To achieve this goal, advanced computer and programming knowledge are essential. Thus, I joined the MSc(CS) programme this year.

It is glad to know HKU has launched a new master programme, Master of Data Science, which covered the latest statistical techniques like machine learning and big data analysis. The programme is also designed in a practical matter to allow students to solidify their knowledge through implementation. It combines the edges of both MStat and MSc(CS) and I would be happy to recommend the others who are interested in quantitative research to study this programme.

WONG Cheuk Lam *CFA, FRM*
Head of Research, Temujin Capital Management

To succeed in data science in the commercial world, you must be excellent in the fields of Computer Science, Statistics and Business Intelligence. That was what I found at the beginning of my career. As a computer engineering graduate, I started my career in Data Science and gradually found there was an obstacle to have advancement on my career. I realized that I should gain more solid skills and knowledge in Statistics to understand the algorithms and perfect my analyses. I had chosen Master of Statistics programme from the University of Hong Kong, which was one of the most important decisions in my life.

Data are always around us. The key is not data. Instead, it's how you handle it and what you get from it. The MStat programme provided the statistics skills I needed to build my career as one of the earliest Hong Kong local data scientists. I believe that the newly launched Master of Data Science from HKU, which provides a balanced combination in Computer Science, Statistics and Business Intelligence, would be a good choice for people who want to excel in Data Science.

CHEUNG Chee Fung, Kloser
Data Science Manager, AXA China Region Insurance Company Limited

The MStat programme provided me serval opportunities to work on real-world data science projects that required multi-disciplinary skills, such as computer programming, statistics, and specific domain knowledge. The knowledge and experience provided by the MStat programme enabled me to develop a much deeper understanding of statistical theories, machine learning algorithms, and many other analytic methods. Having a background in information technology and computer programming from my previous education and work experience also allow me to handle large volumes of data, in particular unstructured data like text or images that are required in many emerging analytic methods. Armed with the knowledge from both statistics and computer technology, many job opportunities welcome me upon my graduation from MStat.

Data science and big data technologies are becoming more and more ubiquitous in all different businesses and industries. I think the new interdisciplinary Master of Data Science programme jointly offered by the Department of Statistics and Actuarial Science and Department of Computer Science is very up-to-date, targeting knowledge and skills that are highly demanded in the job market. The curriculum is very well designed to prepare their students to be successful data scientists in the future.

LUI Siu Man, Carrie
Senior Statistical Officer, Hospital Authority

Big Data in HKU

THE UNIVERSITY OF HONG KONG
FACULTY OF SCIENCE
THE BIG DATA RESEARCH CLUSTER

The Big Data Research Cluster (BDRC) of HKU Faculty of Science, launched in 2016, aims to provide HKU wide leadership and strategies in research, education, and technology transfer in the multidisciplinary field of big data, and thereby generate bigger impacts to HKU and the broader community.

"Unlike traditional data, solution to issues in Big Data will usually call for expertise in different disciplines including but not limited to the owner of the data, IT experts, statisticians and experts in the relevant field." said Professor WK Li, Chair Professor of Department of Statistics and Actuarial Science and Director of BDRC.

The HKU Innovative Data Mining Application Award 2018, sponsored by SAS which provided data mining software to HKU, announced the winners back in June 2018. The winning teams produced impressive projects which showcased their ability to apply data mining techniques to solve real-life challenges.

The team winning first prize used a variety of data mining techniques in "Forecasting Dengue Outbreaks in San Juan and Iquitos", and identified key factors influencing the number of dengue patients. Meanwhile, the team winning second prize went "Into the Wilderness" and produced "An Analytics of Forest Cover Types in the Roosevelt National Forest in the US", trying to suggest the suitable forest cover for different areas. The third winning team wondered "Does the Lyrics Tell Us the Mood of a Song?" with the aim to develop a highly efficient song recommendation system.

Suoxinda Scholarship in Data Science

We are very grateful to Shenzhen Suoxinda Data Technology Company Limited for funding Suoxinda Scholarship in Data Science, a scholarship for outstanding postgraduate students pursuing Data Science studies at the University of Hong Kong. The scholarship aims to nurture data scientists for the banking and financial industries.

Scholarship recipients would be selected from the students entering the Master of Data Science programme on the basis of academic merit and admission interview performance.

Examples of backgrounds of admitted students:

CEO	Senior Analyst Programmer
Director	Business Analytics Manager
Vice-President	Compliance Manager
Associate Director	Software Development Engineer
Head of Global Markets	Quantitative Researcher
Principal Application Analyst	Solution Developer
Senior Consultant	Immigration Officer
Senior Manager	Teacher

Tuition Fees

The full tuition fees for the programme is HK$210,000 for the 2019 intake. The fee shall normally be payable in three instalments over 1.5 years for full-time study or in five instalments over 2.5 years for part-time study. In addition, students are required to pay Caution Money (HK$350), refundable on graduation subject to no claims being made, and Graduation Fee (HK$350).

Minimum Requirements

- Applicants shall hold a Bachelor's degree with Honours or an equivalent qualification;

- Applicants should have taken at least one university or post-secondary certificate course in each of the following three subjects (calculus and algebra, computer programming and introductory statistics) or related areas.

- Applicants shall fulfil the University Entrance Requirements.

Application

Online application can be accessed via
https://www.aal.hku.hk/tpg/

Application Deadline

Main Round: December 14, 2018
Clearing Round: 12 noon, January 31, 2019

Programme Director

Professor WK Li

BSc, MA *York*; PhD *W Ont*
Chair Professor
Department of Statistics & Actuarial Science

Enquiries

Miss Aka Lee
Department of Statistics & Actuarial Science
Tel: 3917 4152   Email: mdasc@hku.hk

Programme Details:
https://saasweb.hku.hk/programme/mdasc-index.php

Support for International Students

https://cedars.hku.hk/
Useful information for students:
https://cedars.hku.hk/publication.php

STAFF LIST

Department of Statistics and Actuarial Science

Dr A Benchimol	BSc *UBA*; MA *UAH*; MPhil, PhD *UC3M*
Dr KC Cheung	BSc(Actu Sci), PhD *HK*; ASA
Dr SKC Cheung	BSc *HK*; MSc *ANU*; PhD *CUHK*
Ms OTK Choi	BSc *UBC*; MSc *Oxford*
Dr YK Chung	BSc, MPhil *CUHK*; PhD *HK*
Professor TWK Fung	BSocSc *HK*; MSc *Lond*; PhD *HK*; DIC
Professor FWH Ho	BSc, MSocSc *HK*
Dr F Jiang	PhD *Rice*
Mrs GM Jing	BSc, MA, DipEd *Syd*
Dr CW Kwan	BSc, PhD *HK*
Dr EKF Lam	BA *St. Thomas*; MA *New Brunswick*; PhD *HK*
Professor K Lam	BA *HK*; PhD *Wisconsin*
Dr D Lee	BSc(Actu Sci), MPhil *HK*; PhD *British Columbia*
Professor SMS Lee	BA, PhD *Cantab*
Mr DKT Leung	BA, MBA *HK*
Dr EAL Li	BSc *HK*, Econ, PhD *Syd*
Dr GD Li	BSc, MSc *Peking*; PhD *HK*
Dr WT Li	BSc *USTC*; PhD *Rutgers*
Professor WK Li	BSc, MA *York*; PhD *W Ont*
Dr ZH Liu	ScD *Harvard*
Dr GCS Lui	MSocSc *Birm*; MPhil *CUHK*; PhD *HK*
Mr PKY Pang	BSc *HK*; MBA *NSW*
Dr C Wang	PhD *NUS*
Dr KP Wat	BSc(Actu Sci), PhD *HK*, FRM
Dr RWL Wong	BSc, MPhil *CUHK*; MA, PhD *Pittsburg*; ASA
Professor SPS Wong	BSc, MPhil *HKU*; PhD *Stanford*
Dr JF Xu	BSc *USTC*; MPhil, PhD *Columbia*
Professor HL Yang	BSc *Inner Mongolia*; MMath *Waterloo*; PhD *Alberta*; ASA
Professor JJF Yao	BSc, MSc, PhD *Paris XI*
Professor GS Yin	BSc *Jilin*; MA *Temple*; MSc, PhD *N Carolina*
Dr PLH Yu	BSc, PhD *HK*
Professor KC Yuen	BSc, MSc, PhD *Calgary*; ASA
Dr AJ Zhang	BSc, MPhil *HKBU*; MSc, PhD *Michigan*
Dr ZQ Zhang	BSc *Nankai*, MSc *E China Normal*; PhD *HK*
Dr K Zhu	BSc *USTC*; PhD *HKUST*

Department of Computer Science

Professor B Cautis	PhD *INRIA & University of Paris South, Orsay*
Dr BMY Chan	MS *UC San Diego*, PhD *HK*
Dr KP Chow	MA, PhD, *UC Santa Barbara*
Dr TW Chim	PhD *HK*
Dr LYK Choi	PhD *HK*
Dr RHY Chung	PhD *HK*
Dr WY Chung	PhD *The University of Arizona*
Professor A Montgomery	MBA, PhD *University of Chicago*
Dr D Schnieders	PhD *HK*
Dr M Sozio	PhD *Sapienza University of Rome*
Professor CL Wang	PhD *USC*
Dr RSW Yiu	PhD *UC Berkeley*
Dr SM Yiu	PhD *HK*