Queen Mary
University of London

Programme Specification (PG)

Awarding body / institution:	Queen Mary University of London
Teaching institution:	Queen Mary University of London
Name of final award and programme title:	MSc in Big Data Science
Name of interim award(s):	PGCert, PGDip
Duration of study / period of registration:	1 year
Queen Mary programme code(s):	H6J7
QAA Benchmark Group:	Computing
FHEQ Level of Award:	Level 7
Programme accredited by:	Chartered Institute for IT (BCS), Institution of Engineering and Technology (IET)
Date Programme Specification approved:	
Responsible School / Institute:	School of Electronic Engineering & Computer Science

Schools / Institutes which will also be involved in teaching part of the programme:

NA

Collaborative institution(s) / organisation(s) involved in delivering the programme:

NA

Programme outline

The Big Data Science movement is transforming how Internet companies and researchers over the world address traditional problems. Big Data refers to the ability of managing and exploiting the massive amounts of unstructured data that is generated continuously by companies, users, devices, and extract key information and insights from it.

Data scientists and engineers are highly skilled professionals who are able to combine state of the art computer science techniques for processing massive amounts of data with modern methods of statistical analysis to extract information and insights from data and create new data-driven services. The job market is currently in shortage of trained professionals with that set of skills, and the demand is expected to increase continuously and significantly over the next years.

The programme leverages the world-leading expertise in research at Queen Mary with our strategic partnership with leading IT sector companies to offer to students a foundational MSc in the field of Data Science. The modules of the MSc programme cover the following aspects:

- Statistical data modeling, data visualization and prediction
- Machine learning techniques for prediction and clustering

Queen Mary
University of London

- Big Data Processing techniques for processing massive amounts of data
- Domain-specific techniques for applying Data Science to different domains: computer vision, social network analysis, natural language processing, risk analysis and decision support, etc.
- Use case-based projects that show the practical application of the skills in real industrial and research scenarios.

Compulsory modules covering Big Data Science foundational topics are complemented by elective modules articulated along two streams that correspond to two industry recognised professional roles:

- The Data Science stream is concerned with building scientific and analytics workflows to extract knowledge from data. This is the default stream.
- The Data Engineering stream is concerned with creating scalable computing environments to ingest, process, manage and serve data.

Students will be offered lectures that explain the core concepts, techniques and tools required for large-scale data analysis. Laboratory sessions and tutorials will put these elements to practice through the execution of use cases extracted from real domains. Students will also undertake a large project where they will demonstrate the application of data science skills in a complex scenario.

The programme is offered by academics from the Networks, Centre for Intelligent Sensing, Risk and Information Management, Computer Vision and Cognitive Science research groups from the School of Electronic Engineering and Computer Science. This is a team of more than 100 researchers (academics, post-docs, research fellows and PhD students), performing world leading research in the fields of Intelligent Sensing, Network Analytics, Big Data Processing platforms, Machine Learning, Pattern Recognition, Social Network Analysis, Multimedia, and Multimedia Indexing.

Aims of the programme

The programme will provide students with cutting edge tools, methods, and techniques for analysing large-scale datasets and building big data and machine learning systems. Graduates will be able to pursue careers as data scientists and data engineers in the industry, as well as initiate research in multiple scientific domains that rely on performing advanced data analysis.

The course will cover the following main topics:

- Statistical data modeling, data visualization and prediction
- Machine learning techniques for cluster detection and classification
- Big data processing techniques for processing massive amounts of data
- Cloud and device deployment
- Domain-specific techniques: computer vision, social network analysis, natural language processing, etc.
- Use case-based projects that show the practical application of the skills in real industrial and research scenarios.

This MSc programme recognises the need for training data scientists and engineers and has been designed to maximise student employability on the data science job market. This is achieved by putting together a programme that is:

- Comprehensive, as it offers an end-to-end perspective of data science products, from inception to deployment
- Up-to-date, where each topic is backed-up by world-leading academics and researchers
- Unique, through its carefully designed programme and modules
- Practical, as it emphasises on developing practical skills supported by a rigorous understanding of the underlying principles

More specifically, this programme aims to:

- Enable students to acquire the essential knowledge, skills, competency and scientific awareness needed for a successful career in data science industries
- Develop systematic awareness of the current development of data science, data engineering and data analytics
- Master topic-specific expertise so that they develop expertise in applying scientific knowledge, mathematics and ingenuity to develop advanced solutions and products for technical, societal and commercial problems
- Equip students with practical and transferable big data and machine learning programming skills

Queen Mary
University of London

What will you be expected to achieve?

Students who successfully complete this programme will be able to understand:
- The design principles underlying modern big data systems
- The design principles underlying modern machine learning systems
- The methodology and tools needed to deploy big data systems
- The methodology and tools needed to deploy machine learning systems
- The methodology and tools needed to continuously operate big data and machine learning systems in production
- Merits of different system deployment and operation options, including edge and cloud options

Academic Content:	
A1	Statistical modeling of real data sources for trend detection and prediction
A2	Programming tools and techniques for processing massive amounts of data such as Map/Reduce and Hadoop
A3	Methods and techniques for automated classification and pattern recognition

Disciplinary Skills - able to:	
B1	Evaluate the scientific, mathematical and software 'tools' relevant to the problem domain of Big Data science
B2	Develop novel techniques for analyzing unstructured data sources
B3	Establish hypotheses on data sources, and validate them through statistical techniques

Attributes:	
C1	Engage critically with knowledge in the domain of Big Data science
C2	Develop a global perspective on the sources and uses of new data
C3	Develop information expertise in the domain

How will you learn?

Each module unit involves lectures, problem solving coursework and practical sessions. Lectures are used to introduce principles and methods and also to illustrate how they can be applied in practice. Coursework allows students to develop their skills in problem-solving and to gain practical experience.

Tutorial sessions actively engage students on applying the techniques and tools presented in the lectures to solve practical problems. These sessions take the form of exercise classes and programming laboratories under the guidance of the teaching and research staff.

Students will have the opportunity to work on a project under the supervision of academic and research staff. Projects will either

Queen Mary
University of London

be significantly development based or have a research focus that will require you to undertake practical work. All projects will be expected either to investigate or to make use of techniques that are at the leading edge.

How will you be assessed?

The assessment of the taught course units takes place through a combination written examination and coursework. Coursework can include a combination of quizzes, lab activities or mini-projects.

The MSc project is examined on the basis of a written report, a formal oral presentation, and a demonstration.

How is the programme structured?

Please specify the structure of the programme diets for all variants of the programme (e.g. full-time, part-time - if applicable). The description should be sufficiently detailed to fully define the structure of the diet.

The course is organised in three semesters:

-The first semester (Semester A) consists of three compulsory modules plus one elective module.
- The second semester (Semester B) has two compulsory module plus two elective modules that are chosen among a set of options.
- Students carry out a large project full-time in the third semester (Semester C). This project will be individually supervised by a academic or research staff and will involve some preliminary work during semesters A and B.

Compulsory modules are:
- Sem A: ECS764P Applied Statistics, ECS7020P Principles of Machine Learning, ECS766P Data Mining
- Sem B: ECS765P Big Data Processing, ECS7026P Neural Networks and Deep Learning

Elective modules are organised in streams and allow students to further develop their chosen professional profile:
- Data Science stream: ECS763P Natural Language Processing (Sem A), ECS757P Digital Media and Social Networks (Sem B) and ECS7005P Risk and Decision Making for Data Science and AI (Sem B)
- Data Engineering stream: ECS781P Cloud Computing (Sem A), ECS796P Distributed Systems (Sem B), ECS7028P Data Semantics (Sem B)

Core module:
- Project

Academic Year of Study

Module Title	Module Code	Credits	Level	Module Selection Status	Academic Year of Study	Semester
Applied Statistics	ECS764P	15	7	Compulsory	1	Semester 1
Data Mining	ECS766P	15	7	Compulsory	1	Semester 1

Queen Mary
University of London

Programme Title: MSc in Big Data Science

Module Title	Module Code	Credits	Level	Module Selection Status	Academic Year of Study	Semester
Principles of Machine Learning	ECS7020P	15	7	Compulsory	1	Semester 1
Cloud Computing	ECS781P	15	7	Elective	1	Semester 1
Natural Language Processing	ECS763P	15	7	Elective	1	Semester 1
Big Data Processing	ECS765P	15	7	Compulsory	1	Semester 2
Neural Networks and Deep Learning	ECS7026P	15	7	Compulsory	1	Semester 2
Risk and Decision-Making for Data Science and AI	ECS7005P	15	7	Elective	1	Semester 2
Digital Media and Social Networks	ECS757P	15	7	Elective	1	Semester 2
Data Semantics	ECS7028P	15	7	Elective	1	Semester 2
Distributed Systems	ECS796P	15	7	Elective	1	Semester 2
MSc Project	ECS750P	60	7	Core	1	Semesters 1-3

What are the entry requirements?

A high (2:2) honours BSc in Computer Science, Electronic Engineering, Maths, Physics or related disciplines is required. International students must have English Language skills to a recognised standard. The minimum requirement is: IELTS 6.5, TOEFL (CBT) 237, 92 (iBT) or TOEFL (written test) 580.

Good knowledge of computer programming is highly recommended for students. The Data Engineering stream should be considered by students who have excellent programming skills and are capable of developing their programming skills independently.

How will the quality of the programme be managed and enhanced? How do we listen to and act on your feedback?

The Staff-Student Liaison Committee (SSLC) provides a formal means of communication and discussion between Schools and its students. The committee consists of student representatives from each year in the school/institute together with appropriate representation from staff within the school/institute. It is designed to respond to the needs of students, as well as act as a forum for discussing programme and module developments. SSLC meet regularly throughout the year.

Each school operates a School Education Committee (EduComm), or equivalent, which advises the School/Institute Director of Taught Programmes on all matters relating to the delivery of taught programmes at school level including monitoring the application of relevant QM policies and reviewing all proposals for module and programme approval and amendment before submission to Taught Programmes Board. Student views are incorporated in this Committee's work in a number of ways, such as through student membership, or consideration of student surveys.

All schools operate an Annual Programme Review of their taught undergraduate and postgraduate provision. The process is

Queen Mary
University of London

normally organised at a School-level basis with the Head of School, or equivalent, responsible for the completion of the school's Annual Programme Reviews. Schools/institutes are required to produce a separate Annual Programme Review for undergraduate programmes and for postgraduate taught programmes using the relevant Undergraduate or Postgraduate Annual Programme Review pro-forma. Students' views are considered in this process through analysis of the NSS and module evaluations.

What academic support is available?

All students will be assigned an advisor with whom they may have regular or arranged meetings. In addition the students will have all the standard induction, advice and supervisory arrangements normally offered to students within EECS.

The school handbook will be provided (and made accessible at all times) to students, where all the channels of support will be outlined. These include the support channels within the school and also those available at College level.

Programme-specific rules and facts

How inclusive is the programme for all students, including those with disabilities?

Queen Mary has a central Disability and Dyslexia Service (DDS) that offers support for all students with disabilities, specific learning difficulties and mental health issues. The DDS supports all Queen Mary students: full-time, part-time, undergraduate, postgraduate, UK and international at all campuses and all sites.

Students can access advice, guidance and support in the following areas:
• Finding out if you have a specific learning difficulty like dyslexia
• Applying for funding through the Disabled Students' Allowance (DSA)
• Arranging DSA assessments of need
• Special arrangements in examinations
• Accessing loaned equipment (e.g. digital recorders)
• Specialist one-to-one "study skills" tuition
• Ensuring access to course materials in alternative formats (e.g. Braille)
• Providing educational support workers (e.g. note-takers, readers, library assistants)
• Mentoring support for students with mental health issues and conditions on the autistic spectrum.

Links with employers, placement opportunities and transferable skills

The staff involved in the MSc of Big Data Science have strong links and research collaboration with industrial partners including IBM, HP, BBC, and Tech City IT startups. Several of these companies will be involved in the teaching activities, providing guest lectures, as well as business use cases for applying Big Data Science techniques.
Additionally, several of the MSc projects offered to the students will be performed in collaboration with an industry partner, including summer placement opportunities.

Programme Specification Approval

Person completing Programme Specification: Dr Jesús Requena Carrión

Queen Mary
University of London

Person responsible for management of programme:

Dr Ahmed Sayed

Date Programme Specification produced / amended by School / Institute Learning and Teaching Committee:

21 Nov 2023

Date Programme Specification approved by Taught Programmes Board: