



## 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:    | Master of Science (MSc) MSc in Big Data Science with Machine Learning Systems |
| Name of interim award(s):                   | N/A   |
| Duration of study / period of registration: | 1 calendar year   |
| Queen Mary programme code(s):               | I4U7  |
| QAA Benchmark Group:                        |   |
| FHEQ Level of Award:                        |   |
| Programme accredited by:                    |   |
| 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:

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

### Programme outline

Data science is transforming how companies, researchers, governments and other organisations over the world address traditional problems. Two main knowledge areas are responsible for the success of the data science movement. First, big data technologies are allowing to exploit massive amounts of unstructured data that are continuously generated by companies, users and devices. Second, machine learning techniques are being developed and deployed to extract understanding from data and create data-driven systems capable of predicting the behaviour of real-world systems.

A data scientist is a highly skilled professional who is able to combine state of the art computer science techniques for processing massive amounts of data with modern methods of statistical analysis to extract understanding from data and create new knowledge-based services. The job market is currently in shortage of trained professionals with this set of skills, and the demand is expected to increase significantly over the next years.

The School of Electronic Engineering and Computer Science at Queen Mary University of London performs world-leading research in data science and since 2014 has offered one of the first of its kind and most popular programmes in this area worldwide, namely MSc in Big Data Science. The new MSc in Big Data Science with Machine Learning Systems focuses on data science career paths concerned with the development, deployment and production of solutions and data-driven services

through machine learning systems. The MSc in Big Data Science with Machine Learning Systems is articulated around modern machine learning, big data technologies and software engineering and teaches how to design, build and maintain end-to-end machine learning production systems, both on-device and served by cloud platforms. This is a new and growing field known as MLOps.

The course leverages our world-leading expertise in research at Queen Mary University of London with our strategic partnership with leading IT sector companies and through its offering covers a wide variety of topics including:

- 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.

Students will be offered lectures that explain the core concepts, techniques and tools required for large-scale data analysis and machine learning development and deployment. 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 for Multimedia Pattern Recognition, Social Network Analysis, and Multimedia Indexing.

## Aims of the programme

The course 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 programme 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 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

## 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:

|     |  |
|-----|--|
| A 1 | Comprehensive knowledge and understanding of scientific principles and methodology necessary to underpin their education in big data and machine learning systems, and an understanding and know-how of the scientific principles of related disciplines, to enable appreciation of the scientific and engineering context, and to support understanding of the relevant historical, current and future developments and technologies. |
| A 2 | Understanding of programming tools and techniques necessary for processing massive amounts of data, including cold processing, stream processing and graph processing.   |
| A 3 | Statistical modeling of real data sources for pattern discovery, trend detection, prediction, and practical understanding of methods and techniques  |
| A 4 | A comprehensive knowledge and understanding of computational models relevant to all machine learning topics on offer, and an appreciation of their limitations and use cases   |

### Disciplinary Skills - able to:

|     |   |
|-----|---|
| B 1 | Evaluate the scientific, mathematical and software tools relevant to the domain of big data science   |
| B 2 | Develop novel techniques for analyzing unstructured data sources  |
| B 3 | Establish hypotheses on data sources, and validate them through statistical techniques  |
| B 4 | Ability to apply and integrate knowledge and understanding of machine learning disciplines, evaluate them critically and apply them effectively   |
| B 5 | Ability to identify, classify and describe the performance of machine learning algorithms through the use of analytical methods and modelling techniques.   |
| B 6 | Ability to formulate problems, identifying constraints including environmental and sustainability limitations; ethical, health, safety, security and risk issues; intellectual property; codes of practice and standards. |

### Attributes:

|     |   |
|-----|---|
| C 1 | Engage critically with knowledge in the domain of Big Data science and Machine Learning Systems |
| C 2 | Develop a global perspective on the sources and uses of new data                                |
| C 3 | Develop information expertise in the domain   |

|     |  |
|-----|--|
| C 4 | Be able to communicate their work to technical and non-technical audiences.  |
| C 5 | Develop research capacity: Work with information that may be incomplete or uncertain, quantify the effect of this on the AI design and, where appropriate, use theory or experimental research to mitigate deficiencies. |

### How will you learn?

Each non-project-based 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 staff. In addition to the project, other modules introduce project working skills.

### How will you be assessed?

The assessment of the taught course units takes place through a written examination and coursework.

The project is examined on the basis of a written report, a formal oral presentation, and a demonstration of the piece of work developed by the student. The projects will have two examiners each, with a third if there is disagreement.

### 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 programme is organised in three semesters:

- The first semester starts in January and corresponds to the Semester B in the academic calendar. It consists of five modules that will cover the foundational techniques and tools to create data-driven solutions, including machine learning, big data processing and cloud computing, as well as knowledge domains that are central in the data industry.
- The second semester corresponds to the short Semester C and is delivered over the summer. It consists of three advanced modules that build on the modules offered during the first semester. Topics covering advanced aspects of machine learning are covered as well as the design, deployment and maintenance of end-to-end machine learning production systems.
- Students carry out a large project full time in the third semester, which corresponds to Semester A and starts in September. Students agree to work on a topic with a supervisor during the first semester, and complete the preparation reading over the second semester.

### Academic Year of Study

Programme Title: MSc in Big Data Science with Machine Learning Systems

| Module Title                      | Module Code | Credits | Level | Module Selection Status | Academic Year of Study | Semester   |
|-----------------------------------|-------------|---------|-------|-------------------------|------------------------|------------|
| Principles of Machine Learning    | ECS7020     | 15      | 7     | Compulsory              | 1                      | Semester 2 |
| Big Data Processing               | ECS765P     | 15      | 7     | Compulsory              | 1                      | Semester 2 |
| Cloud Computing                   | ECS781P     | 15      | 7     | Compulsory              | 1                      | Semester 2 |
| Information Retrieval             | ECS736P     | 15      | 7     | Compulsory              | 1                      | Semester 2 |
| Digital Media and Social Networks | ECS757P     | 15      | 7     | Compulsory              | 1                      | Semester 2 |
| Data Mining                       | ECS766P     | 15      | 7     | Compulsory              | 1                      | Semester 3 |
| Neural Networks and Deep Learning | ECS7XX      | 15      | 7     | Compulsory              | 1                      | Semester 3 |
| Machine Learning Deployment       | ECS7XX      | 15      | 7     | Compulsory              | 1                      | Semester 3 |
| Project                           | ECS751P     | 60      | 7     | Core                    | 1                      | Semester 1 |

### 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.

### 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 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. Staff-Student Liaison Committees meet regularly throughout the year.

Each school operates a Learning and Teaching Committee, 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 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 a tutor, with whom they will have bi-weekly 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 with Machine Learning Systems 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 and machine learning 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

Programme Title: MSc in Big Data Science with Machine Learning Systems

**Person responsible for management of programme:**

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

**Date Programme Specification approved by Taught Programmes Board:**