



PROGRAMME OVERVIEW

The opportunity to exploit Big Data is recognised world-wide and some countries include it in their economic strategies. The UK Government identified Big Data as one of the 8 great technologies which will have a strong impact on growth and the Scottish Government highlights it as an emerging opportunity for Scotland.

Our MSc in Data Science aims to produce specialist data scientists with training in industry relevant data acquisition, storage, warehousing, analytics and visualisation tools and techniques and a good understanding of the needs of industry. The course will prepare graduates in technical disciplines for a career in the design and implementation use of computer-analytics and visualisation solutions for industry.

Contact Dr. Eyad Elyan,
for further information:
01224 262723, e.elyan@rgu.ac.uk

LEARNING

- Group Project
- Individual Project

Throughout the course, content is complemented by practical work, allowing you to support your theoretical knowledge with practical experience in data storage, mining, warehousing, visualisation and analysis as well as transferrable skills. You will be taught through a mixture of lectures, tutorials, labs. You will be invited to attend talks presented by highly-experienced researchers, speakers from industry, and members of the BCS (British Computer Society) on a wide range of industry-related topics. You will also be supported through our online virtual learning environment where you can access a wide variety of resources and other support materials.

The individual project provides an opportunity for applying specialist knowledge together with analytic, problem-solving, managerial and communication skills to a particular area of interest within data science. Working with the full support and guidance of an allocated project supervisor, you will be given the opportunity to propose, plan, specify, develop, evaluate, and present a substantial project.

COURSE CONTENT

Data Warehousing

Data Mining

Data Management

Intelligent Web Technologies

**Professional Skills for Data
Scientists**

**Big Data Analytics and
Visualisation**

**Data Science Development
Advanced Data Science**

MSc Project

BENEFITS

Benefits to You

The course will focus on satisfying industry's demand for data scientists. Upon completion of the course You will be able :

- Apply appropriate data science tools and techniques to industry's data in order to uncover important, previously unknown information only implicit in the data.
- Handle large amounts of real-time, non-persistent, data.
- Contribute to business decision-making by effectively communicating (potentially large volumes of) key data visually.
- Understand, clean up, summarise, interpret and manage data.
- Grasp key knowledge about new problem areas in order to communicate with end-users; understand key business needs and processes and identify added value through data analytics.
- Provide user-centred data analytics at an appropriate level.
- Protect and share data as appropriate.

As part of Robert Gordon University we also have an outstanding employment rate for our postgraduate students at 97%* and are a top-rated university for employability.

OPPORTUNITIES

The course prepares you for a career in Data Science. Job openings include: Data Scientist, Data Analyst, Data Visualisation Specialist, Data Manager, Database Designer/Manager, Data Mining Expert and Big Data Scientist.

Aberdeen is home to many multinational oil and gas companies and associated suppliers such as mainstream software houses, IT providers to major oil-related companies, specialist software consultancies, and venture capital start-ups.

The university is involved in a number of commercial collaborations on a local, national and international scale with organisations such as BP, British Geological Survey, Wood Group PSN, Accenture, WIPRO and many Aberdeen-based software development companies.

The course also prepares students for research careers by providing the skills necessary of an effective researcher. Suitable MSc graduates may continue to PhD programmes within the school. Further information can be found the School of Computing and Digital Media Postgraduate Research pages.

For fees & more information,
please visit: <http://www.rgu.ac.uk/datascience>

WHO SHOULD ATTEND?

The normal minimum entry requirements are an Honours Degree (2nd class or better) in a discipline with significant Computing, Mathematics or Statistics content. All international students, for whom English is not their first language, must provide evidence of linguistic ability, by gaining either IELTS 6.0 or its equivalent prior to receiving an unconditional offer of a place on the course.

KEY DETAILS

- Full/Part-time
- 1 Year full-time, 24 - 30 months part-time
- September start dates
- MSc Final award
- PgCert/PgDip/MSc Exit award
- Course Leader: **Dr. Eyad Elyan**

Funding Available

Visit www.rgu.ac.uk/scholarships
A number of fully free funded places are available for Scottish/EU applicants for 2015/16

MODULES

Data Management (15 Credits)

Data Mining (15 Credits)

Data Warehousing (15 Credits)

Information Retrieval Systems (15 Credits)

Professional Development and Research Skills (15 Credits)

Big Data Analytics and Visualisation (15 Credits)

Data Science Development (15 Credits)

Advanced Data Science (15 Credits)

MSc Project Investigation (15 credits)

MSc Project (45 Credits)

INDICATIVE MODULE CONTENT

Conceptual modelling: E-R modelling. The relational database model: tables, relationships, keys, joins, and normalisation. Physical database design: designing indexes, user views and security mechanisms. Data management standards in industry: data management challenges for industry; big data and data streams; tools, standards and techniques for the management, storage, querying and transfer of data (e.g., PPDM data model, WITSML, PRODML, etc.) Database types: relational, object, noSQL, databases for the internet.

Data mining concepts. Implementation of fundamental learning approaches. Rules involving relations; incorporating domain knowledge in learning. Advanced techniques for evaluating learned concepts. Calculation of confidence intervals for predictive performance. Comparison of data mining schemes. Paired t-test. Minimum Description Length principle

Data Capture, data cleaning, data conformation, data integration, data federation and data virtualisation. Concepts and benefits associated with data warehousing. Conventional, spatial and temporal data warehouses. Architecture of a data warehouse. Data warehouse design. Tools for Data warehousing. State of the art in data warehousing, including data warehousing in the cloud. Data warehousing with big data.

Information collection: crawling and document/content pre processing. Information retrieval: indexing, search, and retrieval. Content: Web content, heterogeneous data, image/video/audio and multimedia content. Web and information environments: mobile information, context aware retrieval, ambient computing, cross/multilingual systems, and social media.

Library skills: literature searches; information sources (on-line and off-line). Legal Issues: Social and ethical implications of IT; copyright; patents; intellectual property rights; contracts; product liability; data protection. Project planning and management: skills, tools and techniques; configuration management.

NoSQL data stores (key-value, document, and graph), e.g. MongoDB. Case studies of NoSQL data stores with hands-on experience. Schema migration in NoSQL data stores (key-value, document, and graph). Modern parallel data processing techniques, e.g. MapReduce/Hadoop. Case studies on using parallel data processing for analysis and mining of Big Data. Visualisation techniques for Big Data.

Data science programming concepts. Data preparation methods. Data exploration, summarisation, transformation and visualisation techniques. Descriptive analytics (Cluster and link analysis). Predictive analytics (Classification and regression analysis). Advanced analytics (Text mining and social network analysis).

Fundamentals of data stream mining: mining strategies and change detection methods. Hoeffding-bound and windowing techniques for data streams: Hoeffding bound as a generic method to data stream mining, and the concept of window-based mining. Stationary data stream mining techniques: mining in a setting where the data source generates a stable distribution of the data. Evolving data stream mining techniques: mining in a setting where concept drift is anticipated. Application of streaming techniques to social media generated data.

Selection of an approved university-based or industry project. Investigation of problem, including context, background, and relevant tools, methods and techniques. Summary of results / research conclusions. Development of a project specification. An ethical review of the project, together with a plan to address any ethical issues.

An overall development plan and breakdown of activities. An overall and a detailed research strategy / design specification. A test/evaluation specification and schedule. A review document with recommendations for future research / development. A practical demonstration of the operational project and conclusions, including an oral presentation and poster. A final report and CD including all software and documentation.

COURSE CALENDAR: SEPTEMBER 2016

Semester One Induction Week

Start of teaching
12 weeks of taught modules

- Data Management
- Data Mining
- Data Warehousing
- Information Retrieval Systems

First semester assessment
Post semester break

Semester Two

Start of Second Semester
12 weeks of taught modules

- Professional Development and Research Skills
- Big Data Analytics and Visualisation
- Data Science Development
- Advanced Data Science

Second Semester Assessments (and Resits of Sem1)
Post assessment break

Week beginning **19 Sep 2016**

27 Sep 2016

7-21 Jan 2017

23-27 Jan 2017

1 Feb 2017

6-20 May 2017

22-26 May 2017

Semester Three

Start of Semester 3

- MSc Project Investigation
- MSc Project

Resit Assessments of End of Sem2
Course Graduation

29 May 2017

5-9 Aug 2017

8 Sep 2017

December 2017