UNIVERSITY OF EXETER

DEGREE APPRENTICESHIPS

High-Level Summary:

Level 7 Research Scientist (MSc Data Science Professional)

Degree Apprenticeship

Programme Overview

The Level 7 Research Scientist (MSc Data Science Professional) programme is an innovative apprenticeship and part-time study programme designed for professionals who wish to gain an MSc and are unable to study full-time. The programme delivers the Level 7 Research Scientist Apprenticeship Standard, against which we deliver a Post-graduate Diploma. For those candidates who wish to top this up to an MSc, there is a final 60 credit Research Project. More information about the Apprenticeship Standard can be found here.

This programme combines the academic rigour of Exeter's long tradition of teaching excellence with the achievement of an industry-recognised professional qualification. The course will cover the core areas of data science (e.g. machine learning, statistics) as well as underpinning tools (e.g. programming, mathematics), specific applications (e.g. network analysis, text analysis, machine vision) and social context (e.g. governance, ethics, business applications), subject to the option module chosen by the apprentice.

Three day face to face teaching blocks delivered on campus at University of Exeter will begin each term. These teaching blocks will then be complemented by a day a week distance learning sessions for the remainder of the term. The weekly distance learning sessions will be comprised of live content delivered virtually alongside independent study.

Assessment where possible will be focused on the organisational context. For those who want to undertake the MSc top-up, there is a substantial research project based in their organisation which will be closely supported by academics with relevant expertise. Industry experts are regularly consulted on course content and will sometimes act as guest lecturers, ensuring relevance to modern business needs.

Data science is a growth area with excellent career development potential. Apprentices will benefit from contact with leading academics and gain a respected qualification alongside their current employment. Employers of apprentices participating in the course will gain valuable knowledge and improve the skills base within their organisation, without losing productivity. University of Exeter is a world-class research active institution which regularly features in UK Top-10 and Global Top-150 rankings.

Programme Structure

The MSc Data Science Professional programme is a 3 year programme of study at National Qualification Framework (NQF) Level 7. It is divided into 3 'Stages', with each Stage normally being equivalent to an academic year. The programme is also divided into units of study called 'modules' which are assigned a number of 'credits'. The three Stages will each consist of 60 credits to give 180 credits in total.

The Level 7 Research Scientist standard will be delivered within the first two years of the programme. Over the first two year period, apprentices will accumulate 120 credits which is equivalent to a Postgraduate Diploma qualification. The 120 credits are achieved via taught modules and guided independent study. Apprenticeship funding is only applicable for the first two years of study. Participants who want to complete the MSc will use the third and final year of the programme to undertake a Research Project which will enable them to achieve the full 180 credits required for a full Masters degree qualification.

Apprentices can achieve a Level 7 Research Scientist award and Post Graduate Diploma of Education over 2 years through:

- One x3 day on campus teaching block each Term (September in Term 1, January in Term 2 and May in Term 3)
- a day a week distance learning in Terms 1, 2 and 3, and off the job programme engagement over the summer period.

Apprentices can achieve a MSc Data Science degree in year 3 through;

- undertaking a research project to apply the knowledge learnt in the first two years

Apprentices are graded on assignments and in class tests.

Approximate term timings are;

- Term 1: September to December;
- Term 2: January to April;
- Term 3: April to June.

The table below details the programme structure for the January 2022 intake. (Please note: all intakes after January 2022 will be in September with the next one being September 2022):

	Jan - April	April - Sept	Sept - Jan
Year One	Introduction to Data Science (15 Credits)	Fundamentals of Data Science (15 Credits)	Learning from Data (15 Credits)
	Professional Practice 1 (15 Credits)		
Year Two	Data in Business and Society (15 Credits)	Option Module* (15 Credits)	Work Based Project (15 Credits)
	Professional Practice 2 (15 Credits)		
Year Three	End Point Assessment (0 Credits)	MSc Data Science Research Project (60 Credits)	

*Option modules available may alter. Current option modules include: Machine Learning, Statistical Modelling, Social Networks and Text Analysis; and Machine Vision. Please note, each option module will require a minimum of 10 people for it to run.

Programme Team

Partnership Development Officer

An employer's main contact will be with the Partnership Development Officer in the Innovation, Impact and Business team. They will lead employers through the engagement of the programme, recruitment of apprentices (if recruitment is required) and will provide support throughout the programme duration.

Programme Director

The Programme Director is in charge of the academic content of the programme and is responsible for organizing the academic team that delivers the programme modules. The Programme Director will help to address employer feedback and will be influential in the development of the programmes academic content.

Academic Mentor

Each apprentice is assigned an Academic Mentor. Apprentices will meet with their Academic Mentor at regular Tripartite Review meetings over their time on the programme to review progress. The Academic Mentor will be able to provide guidance and support for apprentices and will be their first port of call should they have any queries.

Programme Administrator

For queries relating to ESFA processes or non-academic aspects of the programme, employers can contact the Programme Administrator. Any queries relating to invoices or payment should also be sent to the Programme Administrator.

Module Overview

This section should be read in conjunction with the programme structure table above which shows how individual modules fit into the programme as a whole.

Year 1 – compulsory modules

Module	Summary	Syllabus Plan
Introduction to Data Science (Professional) **COMM414** **15 credits**	In this module you will learn about the broad and fast-moving field of data science. You will be introduced to the core competencies and application areas associated with data science, including data handling & visualisation, machine learning, statistical modelling, social network analysis and text mining. You will also explore the ways in which data science is transforming business and society. Practical exercises, individual study and group work will	Topics (with associated exercises and seminar discussions) such as: The Data Revolution Exploring Data with Python Machine Learning & Statistics Data in Society & Business

	consolidate your learning and provide the foundations for later study.	Social Networks & Text Analysis
Fundamentals of Data Science (Professional) **COMM415** **15 credits**	Data science depends on a solid grounding in mathematics and programming. In this module, you will learn essential mathematical techniques and programming skills specific to data analysis, including how to apply the mathematical techniques you have learned as part of computational data analysis procedures. Other computational methods with direct relevance to data science and processing of large datasets will also be included, such as data analysis packages for Python. Overall this module will ensure you have the core skills and background knowledge that underpin many central topics in data science, including machine learning, statistical modelling, network analysis and computer vision.	Topics will be chosen depending on the background and experience of the student cohort, but are likely to include: Aspects of linear algebra (e.g.): Vectors, Matrices, Linear transformations, Eigenvalues and eigenvectors, Positive definite matrices, Singular value decompositions; Principal Component Analysis; Linear Discriminant Analysis Programming for data science in Python (e.g): Tools for handling data (Python: numpy, scipy, matplotlib, pandas), Linear algebra in code, Notebooks/markdown; Other topics may be included as appropriate to the skills and background of the student cohort.
Learning from Data (Professional) **COMM416** **15 credits**	One of the primary aims of data science is to effectively use data to make better decisions. This module will introduce you to machine learning and statistical methods for learning from data. You will learn about the principal learning paradigms from a theoretical point of view and gain practical experience through a series of workshops. Throughout the module, there will be an emphasis on dealing with real data, and you will use, modify and write software to implement learning algorithms. It is often useful to be able to visualise data and you will gain experience of methods of reducing the dimension of large datasets to facilitate visualisation and understanding.	Topics (with associated exercises and seminar discussions): Taxonomy of problems and approaches in machine learning and statistical modelling Data description and pre-processing Probabilistic classification Clustering and dimension reduction Linear and logistic statistical models Model assessment, cross-validation, hypothesis Testing Bayesian learning Linear support vector machines Clustering (hierarchical and partitional) Principal component analysis

Professional Practice 1 - Organisational Structures and Objectives within Data Science **COMM420** **15 credits**	The success of data science within an organisation depends not only upon a good understanding of the technical skills involved in data science and machine learning, but also on the ability of the data scientist to identify projects that offer business value. They must then be able to promote any data science projects to stakeholders, plan a project and negotiate with colleagues to ensure its successful delivery. This module will provide you with the business knowledge and awareness to achieve these goals effectively. This module is one of two professional practice modules and will explore organisational structures and objectives and how you contribute towards meeting these objectives including continued professional development personally and for your organisation.	The module will provide teaching on the following topics: Organisational context for data science Leadership, teamwork and collaboration Problem solving and project planning Project management Effective communication Professional conduct, ethics, regulations and personal responsibility Development of self and others

Year 2 – compulsory modules

Data in Business and Society **COMM412** **15 credits**	Data science is revolutionising many aspects of society, with major impacts on industry, business and the public sector. These rapid changes raise many important issues concerning ethics, privacy and governance. In this module you will learn about the social context of data science and how data is used to inform business practices. You will also consider the ethical issues that can arise from machine learning and analysis of "big data", as well as the legal frameworks and legislation relevant to collection and use of data by organisations. This module provides essential background for any data scientist or data manager whose use of data might affect people or organisations.	The module will cover: How businesses use data to build, understand and report their strategic goals Applying current concepts in data and analytics to real examples Using 'Design Thinking' to create information management systems Understanding the legal, ethical and governance considerations around use and analysis of data in social and business contexts. Specific topics will include: Data projects. Using design thinking techniques to understand organisational problems in data management and scope solutions to these. Workshop on "what are data?". Big data, small data and the challenge of capturing the long tail of research. Group discussions around research project topic.

		Workshop on data storage and archiving. Workshop on data dissemination, curation and Open Data, the limits of automation, and the challenges of making data accessible and re-usable. Presentations on research projects, group discussion of data challenges within different types of businesses with varying customer base. Data protection and legal frameworks for data collection, storage and analysis.
Work Based Project **COMM423** **15 credits**	This module will allow you to apply the knowledge learned in other modules in a significant data science project. A data science project may take years, and not all projects experience a full life cycle, sometimes being abandoned for cost reasons or change of business strategy. The project is designed to demonstrate the application of knowledge and skills as it would in the Data Science occupation. It collates the raw unsynthesised research findings and demonstrates analysis (statistical and scientific), synthesis, evaluation, literature review and produces recommendations for the business to consider. The report is a significant and complex project in itself and thoroughly tests both higher and lower order knowledge and skills. The project report must be based on a real research project carried out in the employer's workplace as part of the apprentice's day to day activities.	Activities related to this module will include: Introduction to the module. Identification of suitable projects from within the student's organisation or based on previous taught content Creation of a proper project proposal for the End Point Assessment Organisation Live project and report writing This will lead directly to the live End Point Assessment.

Professional Practice 2 **Ethical and Regulatory Requirements for Professional Conduct Within Data Science** **COMM421** **15 credits**	In this module, you will critically reflect on the work you do within your employed role and how the learning you have undertaken in other modules has been applied in the workplace. The focus of the module will be on the Knowledge, Skills and Behaviours (KSBs) required for the Level 7 Research Scientist Apprenticeship standard. The KSBs are delivered throughout the MSc Professional Data Science programme, but must be evidenced in the workplace to complete the Apprenticeship. This module will allow you to reflect on the KSBs and develop a portfolio of evidence in the style and format required for the apprenticeship. The aim is to allow integration and internalisation of skills and knowledge. You will be expected to become a reflective practitioner, consciously reflecting on your own practice and adapting accordingly. This module is the second of two professional practice modules and will explore ethical and regulatory requirements for professional conduct.	The module will cover the following themes: Introduction to reflective practice and work-based learning. The Level 7 Research Scientist Apprenticeship standard; associated Knowledge, Skills and Behaviours; and assessment criteria. Developing your portfolio of evidence. Throughout the module there will be an emphasis on practical activities and discussion

Year 2 – Option modules

Machine learning (Professional) **COMM417** **15 credits**	Machine learning has emerged mainly from computer science and artificial intelligence, and draws on methods from a variety of related subjects including statistics, applied mathematics and more specialized fields, such as pattern recognition and neural computation. Applications are, for example, image and speech analysis, medical imaging, bioinformatics and exploratory data analysis in natural science and engineering. This module will provide you with a thorough grounding in the theory and application of machine learning, pattern recognition, classification, categorisation, and concept acquisition	Topics will include: Introductory material: Practical motivation for machine learning, basic ideas of supervised and unsupervised learning, classification, regression. Describing data. Latent descriptions: k-means, maximum likelihood; mixture models; PCA; ICA. Unsupervised learning: Clustering. Supervised models: k-nearest neighbours, linear and non-linear regression, linear discriminant analysis, logistic regression, SVM (Support Vector Machines) and maximum margin classifiers.

		Loss functions and maximum likelihood estimators. Neural networks and deep learning. Evaluation of performance, dataset balance. Ensemble methods: boosting, bagging, decision trees and random forests Metric learning.
Statistical Modelling **COMM418** **15 credits**	In this course we look at the concepts and methods of modern statistics in greater detail. The course will cover various topics in statistical modelling with Bayesian flavor, including generalised linear models, Hierarchical statistical models, Generative and Discriminative models, Hidden Markov models, use of Markov Chain Monte Carlo and Gaussian processes. The module will include practical application of these techniques as well as theoretical underpinnings and model choice.	Topics will include: Basics of Bayesian statistical modelling Generalised Linear Models Markov Chain Monte Carlo Generative and discriminative models Hierarchical statistical modelling Hidden Markov models Introduction to Gaussian Processes
Social Networks and Text Analysis **COMM419** **15 credits**	The rise of the Web has created huge datasets relating to the interaction of users and online content. Much of this content is relational and is best understood using a network perspective (e.g. hyperlinked web pages; users linking to content; users linking to users on social platforms). Much of this content consists of unstructured text (e.g. webpages, blogs, social media posts) that requires computational methods for analysis at scale. In this module you will learn the core principles of social network analysis and computational text analysis, enabling you to gain insight from the rich data available on the Web.	Social network analysis topics will include: What is a network? Describing networks Visualising networks Network models Community detection Centrality Multiplex of Networks Text analysis topics will include: Words, documents, corpora. Bag-of-words, N-grams, feature extraction. Supervised topic modelling Word2Vec Introduction to sentiment analysis.

Machine Vision **COMM413** **15 credits**	How do we recognise objects and people? How can we catch a ball? How do we navigate our way from our desk to the coffee machine, without bumping into each other? These seemingly simple tasks have represented a challenge for AI scientists for decades. Recent developments in machine vision have seen significant improvement in important applications (face detection cameras, body tracking, and autonomous cars).	Topics will include: Image processing: convolution, linear filters, image gradients Feature extraction & matching: edge & corner detection, multi-scale analysis, feature descriptors, feature matching and tracking Geometric Image formation: geometric transformations, pinhole camera and perspective effects Multi-view geometry and structure from motion: 3D reconstruction Recognition in Computer Vision: classical approaches for image classification, and object detection, k-nearest neighbours and support vector machine classifiers. Deep learning for computer vision: neural networks, convolutional neural networks, object detection, semantic and instance segmentation.

Year 3 modules

End Point Assessment **COMM422** **0 credits**	The credits within this module are attributed for participation in the End Point Assessment (EPA) for the Level 7 Research Scientist Apprenticeship which is delivered at University of Exeter using the MSc Professional Data Science programme. The EPA will be undertaken within a 3-month period after completing all taught modules of the MSc Professional Data Science programme. The EPA consists of two discrete assessments, conducted by an End Point Assessment Organisation (EPAO). To pass this module, you must complete both assessments. These assessments are: **Assessment method 1:** Project report, presentation, and questioning (based on work-based project). You need to achieve a grade	Activities related to this module will include completing the two assessments (Assessment 1 and 2 in the Assessment section) required by the End Point Assessment defined in the 'End-point assessment plan for the Level 7 Research Scientist apprenticeship standard'. There are no specific teaching activities for this module and no new content will be delivered. Instead, students on this module will receive tutorial support from Academic Mentors as they go through the EPA assessments.

	"distinction" or "pass" from the EPAO to pass this EPA assessment method. **Assessment method 2:** Professional discussion underpinned by portfolio of evidence. You need to achieve a grade "distinction" or "pass" from the EPAO to pass this EPA assessment method. This 0-credit module is an essential (non-condonable) part of the MSc Professional Data Science programme and you can not complete the MSc programme if you do not pass the module. It is a 0-credit module delivered on a pass/fail basis and will not affect your credit-weighted MSc grade. To pass this module, you must attempt both EPA assessments, but the outcome of the EPA assessments does not affect this module. That is, receiving a "fail" grade from the EPAO does not mean you will fail this module. However, if you do not attempt either of the EPA assessments you will fail the module.	
MSc Data Science Research Project **60 credits**	This module will allow you to apply the knowledge learned in other modules in a significant data science project. The project report must be based on a real research project carried out in the employer's workplace as part of the apprentice's day to day activities The module will develop your project planning, management and implementation skills as well as those in independent learning, presentation and writing. This will also be an opportunity to gain experience in the implementation of a data science project within your organisation and to develop your understanding of the business requirements of such projects.	Activities related to this module will include: Introduction to the module. Identification of suitable projects from within the student's organisation or based on previous taught content. Creation of a project plan and specification with your supervisor team. Implementation of the project Presentation of progress updates to student and supervisor cohort. Meetings with academic supervisor (face-to-face and/or online). Final submission: a report and final presentation.