Setting up MSc conversion courses to meet employers' Machine Learning and Data Science needs

This discussion document outlines areas to consider that are relevant to setting up MSc conversion courses on Data Science and conversion courses on Machine Learning to meet employers' needs, based on the findings from the BCS independent review commissioned by the Office for Artificial Intelligence.

We have based our advice on existing employer led skills frameworks and standards, plus roundtables with a wide range of employers held over the last eighteen months. We have also consulted with the Institute of Coding in developing this document.

The views and advice in the document are provided by BCS on the basis of our own expertise and good judgement.

BCS, October 2019

BCS - The Chartered Institute for IT
First Floor, Block D
North Star House
North Star Avenue
Swindon SN2 1FA

BCS is a registered charity no: 292786

Table of Contents

1 Long term requirements for MSc Machine Learning and MSc Data Science conversion courses

The long term requirements for MSc conversion courses that support the AI sector deal (as part of the Industrial Strategy) are that they will:

1. Equip students with the expertise and skills necessary for Machine Learning and Data Science related professional job roles, which enable them to successfully progress against established professional skills frameworks, and for which there is demonstrable strong demand from employers.
2. Have programmes that successfully integrate professional development with academic excellence, and in particular through the provision of evidence based, extensive AI work-*related* experience that is assessed against professional skills frameworks.
3. Equip graduates with the ethical competencies necessary to significantly contribute towards embedding ethical principles throughout those aspects of the Machine Learning or Data Science product/service lifecycle that their roles may be concerned with.
4. Have effective, impactful, evidence-based student recruitment and retention policies and processes that are highly inclusive, and fully resourced.
5. Provide an extensive, high quality and properly resourced range of professional development, networking, and mentoring opportunities equally accessible to all students whatever their background, ethnicity or gender.
6. Have a cohort of students that is continuously becoming more diverse, especially gender diverse.
7. Have collaborative partnerships with employers with a strong interest in recruiting postgraduates from such courses.
8. Provide progression routes from and partner with universities that have a track record in producing diverse cohorts of graduate students.

2 Developing new Machine MSc conversion courses: building on experience

New AI MSc courses that start in 2020 should have concrete, credible, resourced plans for eventually meeting the above requirements at the fastest pace that can be reasonably achieved.

Based on past history it seems very likely that robust, independent, but also supportive and enabling accreditation of MSc courses on a regular basis will prove to be an important part of the mix of incentives to achieve these requirements in a timely and objective fashion.

Universities that apply for funding should therefore be able to produce:

- a credible strategy for how they will eventually meet the above requirements,
- an extensive, detailed plan for executing their strategy over the shortest time period possible, and

- evidence of being capable of meeting a substantive amount of the above requirements immediately.

As new MSc Machine Learning and MSc Data Science programmes are established, it may be that part of the ongoing process is to facilitate and coordinate good practice across universities, including around synthesising and influencing standards. This could be integrated into the independent accreditation process if such exists, for example.

In parallel to announcing new funding and setting up the competitive process for applying, it would be useful to begin a round of consultations with employers and universities about putting in place the best support mechanisms to ensure the HE sector can rapidly and effectively deliver the course requirements listed above.

3   MSc Programme Bid Development

To develop a high level business case for MSc conversion courses will require specifying:

- Student proposition: this needs to be a succinct, compelling outline of professional career outcomes, tied to credible pathways from undergraduate degrees. Including credible case studies, e.g. the job of Market Research & Business Intelligence Analyst was recently advertised by the NHS Business Services Authority. The job had a salary in the range £37,570 - £43,772 and required AI and Machine Learning skills (see Section 9). Machine Learning and Data Science skills are also relevant to job roles in business analysis, logistics management, management consultancy, financial services management, and social media CRM (e.g. sentiment analysis).
- Employer proposition - the range and level of skills developed during the course, mapped against relevant industry recognised skills framework. E.g. using SFIA, DDAT and IfATE apprenticeship standards (see Sections 5, 6 and 7).
- MSc work-related experience, which must assess competencies as part of the overall award for MSc. Some universities have used the RITTech professional register as a means of demonstrating undergraduate placements meet employer needs, and this model may be suitable for postgraduate courses as well.
- Outlining criteria for embedding Legal, Social, Ethical and Professional issues (LSEPi) throughout curricula, learning outcomes, teaching methods, assessment, and evidencing criteria for accreditation.

4   Conversion to what exactly?

Very roughly in our view a conversion course may be aimed at either upskilling someone with capabilities and expertise to enhance those they have already or equipping them with entirely new capabilities and expertise aimed at a career change.

That rule of thumb implies two different types of MSc conversion courses may be desirable. Each type may have different kinds of work-related experience, educational practice, and professional development. These two types are:

Upskilling –

MSc programmes focusing on Business Skills: These courses produce MSc graduates with the skills needed to develop business requirements for Machine Learning or Data Science enhanced products and services that will be embedded in the business value chain and to manage them after they are deployed. Graduates will be capable of working as part of an interdisciplinary team that includes AI and Data Science expert practitioners who will be tasked with designing and developing the systems against business requirements. A graduate from such a conversion course would understand enough Machine Learning or Data Science principles, concepts, techniques and methods to be able to develop business requirements that are achievable, scalable, modular and extendable with regards to the possible systems that can be developed and then manage these systems against business needs once they are operational.

These courses should be accessible to arts and humanities students with some prior knowledge of quantitative methods, for example graduates of the Q-step Social Sciences with Quantitative Methods undergraduate degree courses[1].

Career change –

MSc programmes focusing on Applications: These courses produce MSc graduates with the intermediate engineering skills to be able to deploy, manage, and maintain Machine Learning or Data Science technologies within the business value chain. Graduates will be capable of working as part of an interdisciplinary team that includes AI and Data Science expert practitioners who will be tasked with designing and developing systems that meet business requirements. A graduate from such a conversion course would understand enough Machine Learning or Data Science principles, concepts, techniques and methods to support the integration of fully developed AI technologies into existing information technology systems by adapting the existing systems to interface with the new technologies.

These courses should be accessible to arts and humanities students with some prior knowledge of quantitative methods and with some prior experience of software engineering practice, for example graduates of the Q-step Social Sciences with Quantitative Methods undergraduate degree courses who may also have software development experience through previous work.

A third possibility is that a conversion course intends to develop graduates fully capable of designing and developing AI systems that meet employer needs. In this case that may be best achieved through combining a foundation year with an existing (non-conversion) MSc AI course.

5   Core Outcome Requirements across MSc Programmes

This section outlines the core requirements for the set of knowledge, understanding and skills that should be the outcome of MSc conversion courses, whether they specialise in

---

Data Science or Machine Learning, and whether they are focussed on Business Skills or Applications.

The core outcome requirements build on employer led skills frameworks and standards from SFIA, DDAT, the employer led draft AI Skills Framework developed by BCS for the Office for AI, the Institute for Apprenticeship and Technical Education (IfATE) apprenticeship standards and the BCS survey of existing MSc programmes. The full list of standards used and their web-links are given in Section 7. The requirements are intended to allow as much flexibility as possible in the design of MSc programmes, focusing on high level guidance without being overly prescriptive.

5.1 Foundational knowledge and understanding

Science

Graduates from the MSc programmes will have extensive knowledge and understanding of the application of the scientific method to decision making in organisations, including the principles, concepts and techniques for:

- systematically identifying all relevant data in an organisation,
- formulating a comprehensive range of plausible hypotheses based on rigorous exploration and experimentation with the data,
- the iterative evaluation and modification of hypotheses to develop those of optimal utility to the decision-making process.

Graduates will also have a comprehensive understanding of the theory and methods of statistical and mathematical modelling that underpin data driven decision making in largescale, heterogeneous organisations, including theories of how data can be biased and modelling techniques that help minimise misunderstanding or misuse of data.

Engineering

Graduates from the MSc programmes will understand the

- principles, concepts, techniques and quantitative methods necessary to synthesise, analyse, interpret and evaluate structured information from diverse datasets that may be incomplete, inconsistent or unstructured.
- core principles and methodologies for the design, implementation and optimisation of algorithmic systems for information gathering, structuring and analysis that support an organisation's business model including techniques for mitigating against the different ways data may be biased, misunderstood, or misused.

Business

Graduates from the MSc programmes will have extensive knowledge and understanding of:

- organisational strategic requirements involved in selecting models, algorithms, tools and technologies for data processing, data storage, and managing database systems.
- the fundamental aspects of data governance, data security, and communications that affect Data Science practices within an organisation, and systematic approaches within that context for adopting Data Science practices that enhance an

organisation's processes, operations, professional and ethical practice and business outputs.

5.2 Core Skills across all Curricula

Ethics

A Masters graduate will be aware of and take into account ethical concerns relating to the design, development, deployment, management and maintenance of data driven products and services, such as for example

- Unfair or prejudiced bias in data or models
- Potential unconscious bias of data practitioners and product development teams
- Appropriate level of transparency in design and development of data models
- The impact of data science and data engineering on restricting or enhancing user autonomy and wellbeing, whether in the workforce, in the customer base or society at large
- The ability of individuals to have appropriate control over their personal data
- Potential unintended, inappropriate or malicious use of data driven products or services

Business

A Masters graduate can:

- use appropriate methods to identify an organisation's processes, data, priorities and its wider values, objectives and strategy
- collaborate effectively with domain experts to help fully identify organisational requirements, checking understanding and testing models and solutions as they are being developed.
- effectively communicate the value, opportunities and limitations of data science and data engineering technologies to a range of audiences with varying technical background.
- support organisational stakeholders develop appropriate policies, processes and practices to prevent unethical issues arising from the design, development, deployment, or management of data science and data engineering products and services
- proactively support their organisation improve the ethnic and gender diversity and inclusivity of the workforce at all levels.

Other science and engineering skills will vary depending on whether the MSc course is focused on Business Skills or Applications.

6   Additional requirements for MSc programmes

A graduate of an MSc programme will ideally in addition to the above core skills also have a significant subset of the following skills at a level necessary to *contribute* as part of a team to:

- using state of the art Machine Learning or Data Science techniques to model and understand a business and its operation
- scoping business priorities for large or complex changes caused by the adoption of Machine Learning or Data Science products and services
- visualising data analysis outputs using the most appropriate medium to tell compelling and actionable stories relevant for business goals
- identifying relevant data hierarchies or taxonomies and methods for properly documenting and adopting them, such as through machine processing, to meet organisational requirements

The particular subset and emphasis of skills developed by an MSc programme will depend on how much the programme focuses on business skills or application development skills, and the emphasis on Machine Learning compared to Data Science.

A graduate from an MSc conversion course specialising in Machine Learning will also:

- understand the principles, concepts and techniques of Machine Learning needed to address the canonical problems summarised in the 2017 Royal Society report[2] on Machine Learning (which are listed in Section 8 at the end of the document), and how they are applicable to the business needs of an organisation
- be able to critically evaluate Machine Learning research findings and synthesise from them potential impacts, opportunities and challenges for an organisation
- be capable of developing Machine Learning research prototypes that meet organisational business requirements in line with industry standards

7 Employer Led Skills Frameworks and Standards used in this document

The following standards have been used to outline desirable skills MSc graduates should achieve by the end of a one year full time course.

From the Institute for Apprenticeship and Technical Education

- Artificial intelligence (AI) data specialist
  https://www.instituteforapprenticeships.org/apprenticeship-standards/artificial-intelligence-ai-data-specialist/
  Note this standard is still in development. The proposal and standard are approved, but the assessment plan is still under development (as of the time of writing, September 2019)
- Digital and technology solutions specialist (integrated degree)
  https://www.instituteforapprenticeships.org/apprenticeship-standards/digital-and-technology-solutions-specialist-integrated-degree/

---

[2] https://royalsociety.org/~/media/policy/projects/machine-learning/publications/machine-learningreport.pdf

- Data Scientist (integrated degree)
  https://www.instituteforapprenticeships.org/apprenticeship-standards/data-scientist-integrated-degree/

From the Digital, Data and Technology Profession Capability Framework skills standards have been used
- Data scientist: skills they need
  https://www.gov.uk/government/publications/data-scientist-skills-they-need/data-scientist-skills-they-need#what-skills-they-need
- Data Engineer: skills they need
  https://www.gov.uk/government/publications/data-engineer-skills-they-need/data-engineer-skills-they-need

From the employer led Skills Framework for the Information Age (SFIA)
- Skill: Solution architecture - Level 5
  https://www.sfia-online.org/en/framework/sfia-7/en/framework/sfia-7/skills/strategy-architecture/technical-strategy-and-planning/solution-architecture
- Skill: Data management - Level 5
  https://www.sfia-online.org/en/framework/sfia-7/skills/strategy-architecture/technical-strategy-and-planning/data-management
- Skill: Business analysis - Level 5
  https://www.sfia-online.org/en/framework/sfia-7/en/framework/sfia-7/skills/business-change/business-change-management/business-analysis
- Skill: Business modelling - Level 5
  https://www.sfia-online.org/en/framework/sfia-7/en/framework/sfia-7/skills/business-change/business-change-management/business-modelling
- Skill: Requirements definition and management - Level 5
  https://www.sfia-online.org/en/framework/sfia-7/en/framework/sfia-7/skills/business-change/business-change-management/requirements-definition-and-management

8 Royal Society Machine Learning Canonical Problems and Approaches

Below are the canonical Machine Learning problems and exemplary approaches for solving them given in the Royal Society report:

Canonical Problem: Classification - To which category does this data point belong?
Approaches: For example, including topics such as Logistic Regression, Support Vector Machines, Neural Networks, Random Forests, and Gaussian Process Classifiers.

Canonical Problem: Regression - Given this input from a dataset, what is the likely value of a particular quantity?
Approaches: For example including topics such as Linear Regression, Neural Networks, and Gaussian Processes.

Canonical Problem: Clustering - Which data points are similar to each other?
Approaches: For example, including topics such as k-means, Gaussian mixtures, and Dirichlet process mixtures.

Canonical Problem: Dimensionality reduction - What are the most significant features of this data and how can these be summarised?
Approaches: For example, including topics such as Principal Components Analysis, Factor Analysis, Multidimensional Scaling, Isomap, and Gaussian Process Latent Variable Models.

Canonical Problem: Semi-supervised learning - How can labelled and unlabelled data be combined?
Approaches: For example, including topics such as probabilistic models, graph-based semi-supervised learning, and transductive Support Vector Machines

Canonical Problem: Reinforcement learning - What actions will most effectively achieve a desired endpoint?
Approaches: For example, including topics such as Q-learning, direct-policy methods, and PILCO[3].

9 Example occupational progression for an MSc graduate

This section includes two example professional roles that an MSc conversion graduate should be able to progress to after successful completion of their course, based on recent job adverts. The first is an example that an MSc graduate specialising in business skills should be able to progress to in the medium term. The second is an example that an MSc graduate specialising in applications should be able to progress to in the long term. They are included to demonstrate the relevance of the skill requirements outlined in this document to actual professional job roles that have recently been advertised.

9.1 Medium term progression for graduate specialising in business skills

The following is taken from an advert for the role of Market Research and Business Intelligence Analyst with the NHS Business Services Authority, which was advertised in the autumn of 2019 with a salary of between £37,570 - £43,772. It is intended to provide an example of the type of professional role an MSc conversion course graduate specialising in business skills should be able to progress to over the medium term after completion of their studies.

The NHS Business Services Authority (NHSBSA) is a Special Health Authority and an Arm's Length Body of the Department of Health and Social Care (DHSC). They provide a range of critical central services to NHS organisations, NHS contractors, patients and the public. NHSBSA is also the main processing facility and centre of excellence for payment, reimbursement, remuneration and reconciliation for NHS Dental Practitioners, Pharmacists, NHS Pensions and other affiliated parties. They are in the process of transforming their services through digitisation, taking advantage of new technologies such as artificial intelligence and machine learning.

---

[3] http://mlg.eng.cam.ac.uk/pilco/

The role of Market Research and Business Intelligence Analyst was advertised to join a team within the Strategy, Business Development and Growth directorate. The directorate plays a key role in the design, development and co-ordination of NHSBSA wide activities.

The job advert outlined that a typical day might include a request to undertake a market share or competitor analysis for a wide variety of current or future potential services, reviewing and responding to consultation documents from other health bodies, or identifying best practice in the areas of customer experience and innovation. The advert specified that an ideal candidate would be resourceful and proactive in seeking out market and business intelligence to support the team in keeping abreast of emerging developments within the wider healthcare landscape. The advert also specified that applicants needed to have a degree level qualification in a subject requiring high levels of data, numeracy, analytics and/or Business/Customer/Market focus.

A graduate with the range of skills and knowledge outlined in this document should be able to demonstrate they are suitable for this type of professional role over the medium term after completion of their course.

9.2   Long term progression for graduate specialising in applications

The following is taken from an advert for the role of Data Engineer with Lloyds Banking Group, which was advertised in the autumn of 2019 with a starting salary in the range of £58,626 - £65,140 a year. It is intended to provide an example of the type of professional role an MSc conversion course graduate specialising in applications should be able to progress to over the long term after completion of their studies.

Lloyds Banking Group is the UK's biggest Retail, Digital and Mobile bank with over 30 million customers. The advertised role was to join the Digital Development and Applied Sciences Team, to support their work of Artificial Intelligence, Data Science and Robotics. The team is helping to build data science solutions around virtual smart agents, network threat detection, price optimisation, customer segmentation, credit decisions and fraud detection.

The advert specified that the role would require

- working in a cross-functional team, collaborating closely with the business, product owners, data scientists, and machine learning engineers to solve business problems
- providing 'data expertise' to help shape the approach to solving problems and contributing new datasets
- developing high quality, reusable data pipelines at scale, ingesting structured and unstructured data into Lloyds' data lake, and feature engineering alongside leading Data Scientists and Machine Learning Engineers
- working in Hadoop, Spark, and Azure environments, using state of the art toolsets, as well as evaluating new technologies to expand Lloyds' capabilities
- good coding/scripting experience (e.g. Python, Java, Scala, C#, SQL etc) developed in a commercial/industry setting
- experience working with relational and non-relational databases (SQL and NoSQL stores) and some exposure to Hadoop technologies
- and included as highly desirable experience of data analysis using Python

The above skills are beyond those that an MSc graduate specialising in applications would have immediately on completing their studies, but they should be capable of progressing to this level of skills in the long term. On completion of their course they should be equipped with a significant range of skills that are relevant to the above bullet points, particularly the first three.