New College of the Humanities

Machine Learning and Data Mining II Course Descriptor

Course Title	Machine Learning and Data Mining II	Faculty	EDGE Innovation Unit (London)
Course code	NCHNAP564	Course Leader	Professor Scott Wildman (interim)
Credit points	15	Teaching Period	This course will typically be delivered over a 6-week period.
FHEQ level	5	Date approved	June 2020
Compulsory/ Optional	Compulsory		
Pre-requisites	None		
Co-requisites	None		

COURSE SUMMARY

This course continues with supervised and unsupervised predictive modelling, data mining, machine-learning concepts and feature engineering. Covers mathematical and computational aspects of learning algorithms, including kernels, time-series data, collaborative filtering, support vector machines, neural networks, Bayesian learning and Monte Carlo methods, multiple regression, and optimization. Uses mathematical proofs and empirical analysis to assess validity and performance of algorithms. Studies additional computational aspects of probability, statistics, and linear algebra that support algorithms. Requires programming in R and Python. Applies concepts to common problem domains, including spam filtering.

COURSE AIMS

- Train learners in advanced machine learning techniques such as neural networks, Bayesian learning and Monte Carlo.
- Train learners in the mathematical foundations of machine learning and data mining methods.
- To allow learners to explore a range of advanced machine learning and data mining techniques and apply them to data science problems.

LEARNING OUTCOMES

On successful completion of the course, learners will be able to:

KNOWLEDGE AND UNDERSTANDING

K1b	Have critical understanding of the mathematical foundations of advanced machine learning and data mining methods.
K2b	Have knowledge and understanding of the challenges associated with predictive model building and deployment and the use of feature engineering.

SUBJECT SPECIFIC SKILLS

S1b	Build predictive machine learning models, for robust data science applications.
S2b	Critically evaluate different machine learning and data mining tools.
S3b	Apply the principles of feature engineering in relation to supervised and unsupervised data using software tools such as Python feature tools.

TRANSFERABLE AND PROFESSIONAL SKILLS

T1b	Critically evaluate different approaches to problem solving.
T2b	Effectively communicate arguments, analyses and conclusions.
T3bi	Develop logical analyses and conceptual thinking.
T3bii	Demonstrate an effective technical proficiency of written English that uses a wide range of literacy skills and vocabulary selected appropriately to communicate to specialist and non-specialist audiences.

TEACHING AND LEARNING

This is an e-learning course, taught throughout the year.

This course can be offered as a standalone short course.

Teaching and learning strategies for this course will include:

- On-line learning
- On-line discussion groups
- On-line assessment

Course information and supplementary materials will be available on the College's Virtual Learning Environment (VLE).

Learners are required to attend and participate in all the formal and timetabled sessions for this course. Learners are also expected to manage their self-directed learning and independent study in support of the course.

The course learning and teaching hours will be structured as follows:

- Off-the-job learning and teaching (6 days x 7 hours) = 42 hours
- On-the-job learning (12 days x 7 hours) = 84 hours (e.g. 2 days per week for 6 weeks)

- Private study (4 hours per week) = 24 hours

Total = 150 hours

Workplace assignments (see below) will be completed as part of on-the-job learning.

ASSESSMENT

FORMATIVE

Learners will be formatively assessed during the course by means of set assignments. These will not count towards the final degree but will provide learners with developmental feedback.

SUMMATIVE

Assessment will be in two forms:

AE	Assessment Type	Weighting	Online submission	Duration	Length
1	Practical skills assessment based on workplace datasets	60%	Yes	Requiring on average 20-30 hours to complete	-
2	Written assignment	40%	Yes	-	1,500 words +/- 10%, excluding data tables

FEEDBACK

Learners will receive formal feedback in a variety of ways: written (via email or VLE correspondence) and indirectly through online discussion groups. Learners will also attend a formal meeting with their Academic Mentor (and for apprentices, including their Line Manager). These bi- or tri-partite reviews will monitor and evaluate the learner's progress.

Feedback is provided on summatively assessed assignments and through generic internal examiners' reports, both of which are posted on the VLE.

INDICATIVE READING

Note: Comprehensive and current reading lists for courses are produced annually in the Course Syllabus or other documentation provided to learners; the indicative reading list provided below is used as part of the approval/modification process only.

BOOKS

- Alpaydin, E., (2014), *Introduction to machine learning*, Cambridge, Massachusetts: MIT Press
- Allison, P., D., (1999), *Multiple regression: a primer*, Thousand Oaks, Calif.; London: Pine Forge Press
- Kuhm, M., and Johnson, K., (2019), *Feature Engineering and Selection: A Practical Approach for Predictive Models*, Chapman and Hall

JOURNALS

Learners are encouraged to consult relevant journals on machine learning and data mining.

ELECTRONIC RESOURCES

Learners are encouraged to consult relevant electronic resources on machine learning and data mining.

INDICATIVE TOPICS

- Neural networks and Bayesian learning
- Multiple regression
- Feature Engineering

Title: NCHNAP564 Machine Learning and Data Mining II **Approved by: Academic Board** **Location: Academic Handbook/Programme specifications and Handbooks/ Undergraduate Apprenticeship Programmes/BSc (Hons) Data Science Programme Specification/Course Descriptors**					
Version number	Date approved	Date published	Owner	Proposed next review date	Modification (As per AQF4) & category number
2.1	May 2022	May 2022	Scott Wildman	September 2026	Category 1: Corrections/clarifications to documents which do not change approved content.
2.0	January 2022	April 2022	Scott Wildman	September 2025	Category 3: Changes to Learning Outcomes
1.0	June 2020	June 2020	Scott Wildman	June 2025	