H8MLF: Machine Learning Fundamentals

Module Code:	H8MLF
Long Title	Machine Learning Fundamentals APPROVED
Title	Machine Learning Fundamentals
Module Level:	LEVEL 8
EQF Level:	6
EHEA Level:	First Cycle
Credits:	5
Module Coordinator:	
Module Author:	Isabel O'Connor
Departments:	School of Computing
Specifications of the qualifications and experience required of staff	MSc and/or PhD degree in computer science or cognate discipline. May have industry experience also.

Learning Outcomes

On successful completion of this module the learner will be able to:

#	Learning Outcome Description
LO1	Recognize the ethical implications of machine learning
LO2	Apply appropriate data sourcing and handling principles
LO3	Build and evaluate advanced machine learning models in various problem domains
LO4	Extract, interpret and evaluate information and knowledge from non-trivial real-world data sets
LO5	Comprehend, analyze and evaluate key concepts in machine learning

Dependencies

Module Recommendations

No recommendations listed

Co-requisite Modules

No Co-requisite modules listed

Entry requirements	See section 4.2 Entry procedures and criteria for the programme including procedures recognition of prior learning.

H8MLF: Machine Learning Fundamentals

Module Content & Assessment

Indicative Content
Data Mining Methodologies and Ethics in Machine Learning KDD, CRISP-DM, SEMMA. Ethics in data sourcing & handling. Regulatory & Privacy Components (including Data Protection Act). Ethical implications of machine learning
Data pre-processing and transformation (I) Identifying and Handling Missing Values. Handling Outliers. Dimensionality Reduction (PCA, MCA, etc.)
Factors Affecting a Machine Learning Model (I) Bias-Variance Trade-off. Curse of Dimensionality
Factors Affecting a Machine Learning Model (II) Understanding Factors that can affect model performance; e.g. Type III errors, selection bias, measurement errors, improper variable encoding. Ethically assessing biases
Regression (I) What is a Regression Problem?. Simple Linear Regression
Regression (II) Multiple Linear Regression. Linear Model Selection and Regularization
Data pre-processing and transformation (II) Measuring Predictor Importance. Feature Engineering. Understanding, Detecting and Handling (massive) class imbalance
Classification (I) What is a Classification Problem?. Logistic Regression
Classification (II) K-Nearest Neighbours (kNN). Naïve Bayes (NB)
Decision Trees Decision Trees. Appropriate Use Cases. Measuring Node Purity. Pruning
Ensembles Random Forest. Bagging and Boosting Methods (e.g. XGBoost)
Clustering Notions of distance and similarity. Clustering methods: k-means, k-medoids, hierarchical. Cluster evaluation measures: DBIndex, WSSSE, scree plots

Assessment Breakdown	%
Coursework	60.00%
End of Module Assessment	40.00%

Assessments

Full Time

Coursework

Assessment Type:	Formative Assessment	**% of total:**	Non-Marked
Assessment Date:	n/a	**Outcome addressed:**	1,2,3,4,5
Non-Marked:	Yes		

Assessment Description:
Formative assessment will be provided on the in-class individual or group activities.

Assessment Type:	Project	**% of total:**	60
Assessment Date:	n/a	**Outcome addressed:**	1,2,3,4
Non-Marked:	No		

Assessment Description:
Project focusing on the practical application of data processing and machine learning techniques to data sets in order to extract insights and perform predictive analytics. Component parts of this project may be assessed at different dates.

End of Module Assessment

Assessment Type:	Terminal Exam	**% of total:**	40
Assessment Date:	End-of-Semester	**Outcome addressed:**	5
Non-Marked:	No		

Assessment Description:
The end of semester examination will contain essay-style questions examining the theory behind machine learning techniques covered during the semester, and may require some calculation. Marks will be awarded based on clarity, structure, relevant examples, depth of topic knowledge and an understanding of the potential and limits of solutions.

No Workplace Assessment

Reassessment Requirement

Repeat examination
Reassessment of this module will consist of a repeat examination. It is possible that there will also be a requirement to be reassessed in a coursework element.

Reassessment Description
Repeat examination Reassessment of this module will consist of a repeat examination. It is possible that there will also be a requirement to be reassessed in a coursework element.

H8MLF: Machine Learning Fundamentals

Module Workload

Module Target Workload Hours 0 Hours

Workload: Full Time

Workload Type	*Workload Description*	*Hours*	*Frequency*	*Average Weekly Learner Workload*
Lecture	Classroom & Demonstrations (hours)	24	Per Semester	2.00
Tutorial	Other hours (Practical/Tutorial)	24	Per Semester	2.00
Independent Learning	Independent learning (hours)	77	Per Semester	6.42
			Total Weekly Contact Hours	4.00

Module Resources

Recommended Book Resources

Brett Lantz. (2019), Machine Learning with R - Third Edition, Packt Publishing, p.458, [ISBN: 9781788295864].

Gareth James,Daniela Witten,Trevor Hastie,Robert Tibshirani. (2014), An Introduction to Statistical Learning, Springer, p.426, [ISBN: 9781461471370].

Christian Heumann,Michael Schomaker,Shalabh. (2017), Introduction to Statistics and Data Analysis, Springer, p.456, [ISBN: 978-3-319-46162-5].

Supplementary Book Resources

Kartik Hosanagar. (2019), A Human's Guide to Machine Intelligence, Penguin, p.272, [ISBN: 9780525560890].

This module does not have any article/paper resources

This module does not have any other resources

Discussion Note:	