Module Code	Pre-requisite Module codes	Co-Requisite Modules code(s)	ISCED Code	Subject Code	ECTS Credits	NFQ Level (CPD)#
TBD					5	8
Module Title	Machine Learning for Predictive Analytics					

Machine Learning for Predictive Analytics

School Responsible:	School of Computing

Module Overview:

The module aim is to provide the student with an understanding of machine learning techniques and data analytics. The module begins by introducing the student to the core-concepts in machine learning, it then reviews for different families of machine learning algorithms (information based, similarity based, probability based and error based algorithms). The module concludes with a review of the best practice in designing evaluation experiments for machine learning. Throughout the module there is a emphasis on grounding the module content in real world examples and linking the material covered to real world problems. At the end of the module the student will have an excellent foundation in the fields of machine learning and data analytics.

Learning Outcomes (LO):	
On Completion of this module, the learner will be able to	
1	Demonstrate an understanding of the core concepts of machine learning and the major trade-offs and pitfall in machine learning
2	Demonstrate an understanding of the core concepts and standard algorithms used to build predictive models across a range of machine learning approaches, including: information based, similarity based, probability based, and error based methods.
3	Demonstrate and understanding of best practice in evaluating machine learning approaches
4	Design and implement machine learning-based solutions to problems

Indicative Syllabus:

- Introduction to Machine Learning: define the machine learning task of learning from data, distinguish between supervised and unsupervised machine learning, explain how machine learning is an ill-posed problem and the need for inductive bias if learning is to take place, explain the problems that can go wrong with machine learning (underfitting and overfitting), introduced the bias-variance trade-off, the no-free lunch principle, and the need to generalise beyond the given dataset.
- Introduce Supervised Learning: the curse of dimensionality and feature selection, the stationarity assumption and i.i.d data, distinguish between lazy versus eager learners,

Module Code	Pre-requisite Module codes	Co-Requisite Modules code(s)	ISCED Code	Subject Code	ECTS Credits	NFQ Level (CPD)#
TBD					5	8
Module Title	Machine Learning for Predictive Analytics					

distinguish between classification and regression tasks, overview approaches to learning: Probability, Reducing Error, Analogy, Information

- Learning using information (entropy based approaches): review information theory and entropy and show how these concepts can be used to develop decision tree classification models. Explain how decision trees can be used for regression problems. Explain how tree-pruning can improve models generalizability, and explain techniques for tree pruning. Introduce model ensembles in terms of boosting and bagging techniques and Random Forest models.
- Similarity based learning: Introduce the concepts of a feature space and nearest-neighbour algorithms. Review a range of indexes for measuring similarity between data points (e.g., Miknowski distances, cosine similarity, russell-rao, sokal-michener, mahalanobis, inter alia). Introduce different methods for making similarity based models robust to noise (k-nearest neighbour models, weighted k-NN).
- Probability based learning: Review basic probability including marginalization, conditional independence and Bayes' Theorem. Introduce the Naive Bayes classifier and the Maximum A Posterior classification criterion. Review methods for handling continuous features within a probabilistic framework (e.g., binning and probability density functions). Introduce Bayesian Networks models and Markov Chain Monte Carlo techniques.
- Error based learning: Review a range of error functions, including the sum of squared errors. Introduce the method of least squares and the gradient descent algorithm for fitting a multivariate linear model to a dataset. Demonstrate how to train a logistic regression model for binary classification and how this model can be extended to multi-nominal classification. Review how linear models can learn non-linear relationships through the use of basic functions and introduce the Support-Vector Machine approach to classification.
- Introduce the best practice in model evaluation, including the problem of peeking and the design of evaluation experiments (e.g., hold-out test set, cross-validation methods). Review a number of evaluation metrics and match these metrics to characteristics within a dataset. Explain the problem of concept-drift and review methods for evaluating models after deployment.

Learning and Teaching Methods:

The course delivery involves a combination of lectures and labs which may incorporate the use of blended learning techniques as appropriate throughout the delivery.

Total Teaching Contact Hours	39
Total Self-Directed Learning Hours	61

Module Delivery Duration:

This module is delivered over one semester

Module Code	Pre-requisite Module codes	Co-Requisite Modules code(s)	ISCED Code	Subject Code	ECTS Credits	NFQ Level (CPD)#
TBD					5	8
Module Title	Machine Learning for Predictive Analytics					

Assessment		
Assessment Type	**Weighting (%)**	**LO Assessment (No.)**
Final Exam	70	1-4
In class examination	30	1-4
Module Specific Assessment Arrangements (if applicable)		
(a) Derogations from General Assessment Regulations		
(b) Module Assessment Thresholds		
(c) Special Repeat Assessment Arrangements		

Essential Reading: (author, date, title, publisher)

John D. Kelleher, Brian Mac Namee, Aoife D'Arcy, 2015, Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies. MIT Press

Version No:		**Amended By**	
Commencement Date		**Associated Programme Codes**	

Modules that are to be offered as Stand-Alone CPD Programmes must have an NFQ level assigned

*Details of the assessment schedule should be contained in the student handbook for the programme stage.

Date of Academic Council approval ………………………….

Module Code	**Pre-requisite Module codes**	**Co-Requisite Modules code(s)**	**ISCED Code**	**Subject Code**	**ECTS Credits**	**NFQ Level (CPD)#**
TBD					5	8
Module Title	Machine Learning for Predictive Analytics					