IB9PB-15 Machine Learning and Data Analytics

26/27

Department
Warwick Business School
Level
Taught Postgraduate Level
Module leader
Neil Stewart
Credit value
15
Module duration
10 weeks
Assessment
100% coursework
Study location
University of Warwick main campus, Coventry

Description

Introductory description

This module should provide you with the skills you need to design and manage a data science project, and utilize techniques from machine learning. Big Data Analytics is about taking the myriad of data created by all of us, and fashioning it into actionable information.

Module aims

The module will provide hands on experience to acquaint non-specialists with the current machine learning techniques in regression, classification and unsupervised learning using the R programming language.

Outline syllabus

This is an indicative module outline only to give an indication of the sort of topics that may be covered. Actual sessions held may differ.

Sessions 1 & 2

- *base* R and RStudio for reproducible research.
- Introduce some key parts of the `tidyverse` including `dplyr` and `ggplot2`.
- `data.tables`

Sessions 3 & 4 -- Regression

- Revision of regression
- Regression and model trees
- Penalised regression

Sessions 5 & 6 -- Classification

- k nearest neighbours
- decision trees and random forests
- support vector machines
- linear discriminant analysis.

Sessions 7 & 8 -- unsupervised learning

- k-means
- hierarchical agglomerative clustering
- natural language processing and latent Dirichlet analysis

Sessions 9 & 10 -- neural networks

- introducing production-ready systems H2O and TensorFlow for machine learning at scale.
- network analytics

Learning outcomes

By the end of the module, students should be able to:

- Understand research methods in big data analytics and computational social science
- Demonstrate understanding of results of research in big data analytics and computational social science
- Understand and identify links between big data resources and real world events
- Apply methods in data analytics and computational social science

Indicative reading list

Reading lists can be found in Talis

Subject specific skills

1. Apply methods in data analytics and computational social science
2. Perform example analyses
3. Understand research methods in big data analytics and computational social science

4. Demonstrate understanding of results of research in big data analytics and computational social science
5. Understand and identify links between big data resources and real world events

Transferable skills

1. Critically evaluate empirical research
2. Preprocess data sets to allow their subsequent application to real world problems
3. Visualise extensive data sets, applying methods which
   both allow the visualisation consumer to ask their own questions, and methods which directly answer specific questions
4. Formulate hypotheses and apply statistical methods for their evaluation
5. Demonstrate business relevant data analytics skills
6. Write in an academically appropriate way
7. Demonstrate confidence in discussing research results and its practical relevance in the real world

---

Study

Study time

Type	Required
Seminars	10 sessions of 3 hours (38%)
Private study	48 hours (62%)
Total	78 hours

Private study description

self study and reflective learning

Costs

No further costs have been identified for this module.

Assessment

You must pass all assessment components to pass the module.

Assessment group A

	Weighting	Study time	Eligible for self-certification
Assessment component			
Individual assignment	100%	72 hours	Yes (extension)
Reassessment component is the same			

Feedback on assessment

module leader feedback

Availability

There is currently no information about the courses for which this module is core or optional.