

EC9D7-15 Machine Learning and Big Data in Economics

25/26

Department

Economics

Level

Taught Postgraduate Level

Module leader

Nathan Canen

Credit value

15

Module duration

9 weeks

Assessment

100% exam

Study location

University of Warwick main campus, Coventry

Description

Introductory description

Analyses in all fields of Economics nowadays make frequent use of large and detailed datasets ("big data"). Their increased availability opens up many opportunities for applied research, as well as new challenges on how to handle, process, and extract meaningful conclusions from the data. This module provides an overview of recent developments in econometric methods tailored to handle such large datasets, both in supervised and unsupervised methods, including machine-learning techniques. Further discussion is provided on the applicability of those methods relative to other Econometric methods.

Module aims

The aim of the module is to introduce students to the analytical tools and the knowledge to study economic problems using modern data science methods. The module covers up-to-date econometric techniques in big data and machine learning, as well as the challenge posed by identification of parameters of interest. The aim is to present the econometric techniques along with the hands-on implementation in the computer language R. The module suggests a number of interesting applications in Economics.

By the end of the module, students should feel comfortable with implementing both supervised and unsupervised machine learning techniques for economic problems using state-of-the-art computational tools. They should be aware of the interpretation of each technique, their statistical properties and the assumptions required for the latter.

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.

- Principal Components
- Lasso, Adaptive Lasso, Elastic Net, Penalized Logistic Regression
- Random Forest, Regression trees
- Neural Networks
- Topic modelling, text analysis

Learning outcomes

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

- Subject Knowledge and Understanding: Be able to use a variety of modern data-science methods to solve economic questions. The teaching and learning methods that enable students to achieve this learning outcome are: Lectures, seminars, independent study. The summative assessment methods that measure the achievement of this learning outcome are: Exam
- Subject Knowledge and Understanding: Be able to use R to process data and apply data-science methods. The teaching and learning methods that enable students to achieve this learning outcome are: Lectures, seminars, independent study. The summative assessment methods that measure the achievement of this learning outcome are: Exam
- Subject Knowledge and Understanding: Understand under which conditions each method applies and be able to adapt their strategy to the problem studied. The teaching and learning methods that enable students to achieve this learning outcome are: Lectures, seminars, independent study. The summative assessment methods that measure the achievement of this learning outcome are: Exam
- Subject Knowledge and Understanding: Be able to use methods for both predictive and causal purposes. The teaching and learning methods that enable students to achieve this learning outcome are: Lectures, seminars, independent study. The summative assessment methods that measure the achievement of this learning outcome are: Exam
- Subject Knowledge and Understanding: Be able to understand, distinguish, and communicate the differences between correlational and causal analysis in the context of big data and machine learning methods. The teaching and learning methods that enable students to achieve this learning outcome are: Lectures, seminars, independent study. The summative assessment methods that measure the achievement of this learning outcome are: Exam

Indicative reading list

[Reading lists can be found in Talis](#)

Interdisciplinary

Module covers contents that are relevant to also disciplines including Computer Sciences and Statistics. Applications of the method might involve other disciplines.

International

The methods presented in the module may be applied to various contexts; in fact, students are encouraged to bring applications and research questions that are pertinent to their international backgrounds.

Subject specific skills

- Develop a critical understanding of statistical methods for large datasets, including both supervised and unsupervised methods.
- Be able to apply such methods in the open-source language R
- Enhance the capacity to conduct economic analyses autonomously
- Be able to understand, distinguish, and communicate the differences between correlational and causal analysis in the context of big data and machine learning methods

Transferable skills

- Develop a critical understanding of statistical methods for large datasets, including both supervised and unsupervised methods.
- Be able to apply such methods in the open-source language R
- Enhance the capacity to conduct economic analyses autonomously
- Be able to understand, distinguish, and communicate the differences between correlational and causal analysis in the context of big data and machine learning methods

Study

Study time

Type	Required
Lectures	9 sessions of 2 hours (12%)
Seminars	4 sessions of 1 hour (3%)
Other activity	1 hour (1%)
Private study	127 hours (85%)
Total	150 hours

Private study description

Private study to consolidate learning materials.

Other activity description

1 hr revision session

Costs

No further costs have been identified for this module.

Assessment

You do not need to pass all assessment components to pass the module.

Assessment group B

Assessment component	Weighting	Study time	Eligible for self-certification
Centrally-timetabled examination (On-campus) Final module exam with 100% of the marks.	100%		No

- Answerbook Pink (12 page)
- Students may use a calculator

Reassessment component is the same

Feedback on assessment

Feedback on the exam will be given following marking.

[Past exam papers for EC9D7](#)

Availability

Courses

This module is Core for:

- TECS-L111 Postgraduate Taught Economics and Data Science
 - Year 1 of L111 Economics and Data Science
 - Year 1 of L111 Economics and Data Science