Imperial College London

Module Specification

Basic details

		Earliest cohort	Latest cohort
UID		Cohorts covered	2025-26	

Long title: Applied Machine Learning

New code: PHYS70022 | New short title:

Brief description of module *(approx. 600 chars.)*:

This core module will provide the hands on experience of techniques required to analyse large data sets. The course will be taught in the Python computing language and will use standard packages such as numpy, scipy, matplotlib, pandas, Scikit-Learn, Keras and Tensorflow. The course assumes no prior knowledge of Python or any of these packages. You will learn how to implement the different techniques required to analyse data (statistical techniques and machine learning techniques) through working through examples and then analysing different data sets.

558 characters

Available as a standalone module/ short course? N

Statutory details

	ECTS	CATS	Non-credit		
Credit value	10	20	N	HECOS codes	

FHEQ level: Level 7

Allocation of study hours

	Hours	
Lectures	0	
Group teaching	90	*Incl. seminars, tutorials, problem classes.*
Lab/ practical	0	
Other scheduled	0	*Incl. project supervision, fieldwork, external visits.*
Independent study	160	*Incl. wider reading/ practice, follow-up work, completion of assessments, revisi*
Placement	0	*Incl. work-based learning and study that occurs overseas.*
Total hours	250	
ECTS ratio	25.00	

Project/placement activity

Is placement activity allowed? No

Module delivery

Delivery mode	Taught/ Campus	Other	
Delivery pattern		Other	
Delivery term	Term 1	Other	

Ownership

Primary department: Physics

Additional:

teaching departments	
Delivery campus	South Kensington

Collaborative delivery

Collaborative delivery?	N
External institution	N/A
External department	N/A
External campus	N/A

Associated staff

Role	CID	Given name	Surname
Module Leader		David	Colling
Lecturer		Patrick	Dunne
Lecturer		Tim	Evans

Learning and teaching

Module description

Learning outcomes	By the end of the course you will have an understanding and experience of the handling and manipulation of large datasets in Python; you will have able to choose from a range of commonly used machine learning tools and apply your skills with them to different classes of problems; you will be able to justify which tools are appropriate for which classes of problem by assessing both their applicability and accessibility within commonly used software frameworks. This is a broad course during which you will be introduced to a wide variety of problems and techniques. This course is designed to prepare you for your research project for which you will (most likely) concentrate on fewer techniques in more detail.
Module content	The module consists of different of using packages commonly used in the physical sciences (numpy, scipy, matplotlib, pandas, Scikit-Learn, Keras and Tensorflow etc) to analyse different dataset and to understand different classes of problems. The module will also cover topics specific to Big Data applications such as data compression and data curation/processing architectures.
Learning and Teaching Approach	The course will be taught by working through practical examples and carrying out analysis on different datasets, using different machine-learning based methods. These sessions will be delivered in computing labs, during which you will have the chance to interact with the lab demonstrators and with your peers to understand the material.
Assessment Strategy	The course will be assessed entirely through coursework. You will be assigned one "mini-project" every few weeks (with a total of 4), for which you will need to use knowledge gained from this and the statistics core module to complete. These will get progressively more challenging throughout the course. For each mini-project, you will write a short (less than 1000 words) report on your findings, which be formally assessed and used to track your progress in understanding the taught material. These reports, along with the results and relevant codes will be submitted and assessed via Jupyter Notebooks.
Feedback	Feedback on the assessments will be returned to you within a 3 weeks after you submit your mini-project report. Additional feedback will be given by demonstrators during practical sessions, during which the demonstrators will go over model solutions/methods for the mini-project. You will be able to discuss your solutions and how they can be improved with the demonstrators and your peers. These sessions can also be used to work on the mini-projects and the demonstrators will be able to help you with the mini-project assignments.
Reading list	See "Practical Data Analysis and Machine Learning in the Physical Sciences" at

https://imperial.leganto.exlibrisgroup.com/leganto/reading list/lists/39788448460001591
(Accessible before start of course)

Quality assurance

Date of first approval
Date of last revision
Date of this approval

Module leader: David Colling

Office use only

QA Lead
Department staff
Date of collection

Date exported
Date imported

Notes/ comments

Template version 16/06/2017