Analytics, Data Science & Decision Making Summer School 2023

Course List

When selecting your courses, please pay attention to the following:

- The Programme is colour-coded into focuses. Whilst you technically can choose to attend courses from different focuses, we strongly encourage you to stay with one focus. Past delegates have explained that they found it difficult to switch focus mid-week, as the courses build in difficulty, and you get more out of a complete focus.
- Each course is also assigned a level of 'introductory' or 'intermediate'. We do not provide advanced courses, as our delegates should feel encouraged to try focuses with basic experience and enthusiasm.
- Structure: some courses are mostly theoretical, but all of our courses have a practical element to them so you can practice your newly acquired knowledge.
- Duration: some courses run for multiple days up to a maximum of 4 days. In order to understand the course fully, it is recommended that you attended all the days that the course is running. You cannot attend multiple courses on one day.
- Pre-requisites: some courses provide guidance on prerequisites, and knowledge that might be required or useful to have in order to understand the course fully. Please take a look at these.
- Preparation: some courses also require you prepare in advance, or bring a laptop, possibly with certain software installed. If so, you will be given access to these instructions and materials via a shared drive one month before the Summer School begins.

Please note *that whilst we expect all of our courses to run smoothly, there may be alterations or last minute changes beyond our control if a course provider does not deliver as planned. The IADS Summer School will inform you of any changes that may occur as soon as possible. If you enrol on a course that no longer continues, all efforts will be made to transfer you to another course of your choice. No refunds will be given on such occasions.*

Course title: Introduction to Machine Learning
Provider: Dr Michael Fairbank, University of Essex
Date: 24 July 2023 (1 day)
Focus: Machine Learning
Level: Introductory level
Course abstract: The aim of this course is to provide an introduction to Machine Learning and a discussion of the types of problems it is suitable for. The course will then introduce Kernel Machines and show how they can provide robust but flexible classifiers when the number of training points is limited. Course objectives: • Understand how ML combines information from data and prior knowledge • Familiarise with supervised, unsupervised and reinforcement learning problems • Understand loss functions: the 0-1 loss for classification and its shortcomings • Understand how kernel classifiers based on the hinge loss can provide robust yet flexible classifiers.
Practical elements: This course is half theory and half practical. For the practical exercises, you will learn how to implement machine learning pipelines using Python and scikit-learn.
Prerequisites/knowledge: A background in computer science, maths, business and social science with some mathematical skills would be useful. This course requires high-school level maths.
Will I need to bring a laptop? No, delegates will have access to the University's computers. Delegate may also bring their own laptops. If bringing a laptop, you will need to download and install software in advance.
Will I need to prepare anything in advance? No

Course title: Ethics, Data Protection, Security, and Liability in the Age of Big Data
Provider: Dr Audrey Guinchard and Dr Anirban Chowdhury, University of Essex
Date: 24 July 2023 (1 day)
Focus: Big Data
Level: Introductory level
Course abstract: The course will begin with introducing the ethical issues arising from Big Data (with both speakers), before presenting the correlative legal issues that may arise in light of the current EU (GDPR), Brexit, and UK data protection regimes and of criminal law (Dr Audrey Guinchard). Torts and contracts will not be covered.
Practical elements: Starting with real-life examples of ethical issues in the first workshop, the 2 law workshops will build on these use cases to unravel the legal framework and give tools to make sense of the GDPR requirements in a possible post-Brexit world.
Prerequisites/knowledge: None
Will I need to bring a laptop, or download anything in advance? No
Will I need to prepare anything in advance? Review the slides which will be shared beforehand.

Course title: Introduction to Python
Provider: Dr Sefki Kolozali and Mohsin Ali, University of Essex
Date: 24 July 2023 (1 day)
Focus: Introductory Course
Level: Introductory
Course abstract: This Introduction to Python course is for beginners. We aim to introduce fundamental programming concepts using Google Colab. We will introduce variables, data types, casting, string, Booleans, operators, lists, tuples, loops, conditions, functions, and a bit of NumPy. This course is designed for those who are coming from a non-technical background and willing to learn Python to great in summer school.
Prerequisites/knowledge: None
Will I need to bring a laptop, or download anything in advance? No
Will I need to prepare anything in advance? Participants need to have a google account to work on Colab.

Course title: Functions, Control flow, and Automation
Provider: Dr Ben Skinner and Dr David Clark, University of Essex
Date: 24 July 2023 (1 day)
Focus: R
Full details TBC

Course title: A hands-on introduction to non-invasive neural interfaces
Provider: Dr Ian Daly, Dr Sebastian Halder and Prof Reinhold Scherer, University of Essex
Date: Dates TBC (4 days)
Focus: BCI
Level: Introductory level / Intermediate level
Course abstract: Neural interfaces (NIs) are emerging technologies that enable direct communication with the nervous system and provide access to motor, sensory, emotional, and cognitive function, opening up entirely new and unprecedented possibilities for human interaction, neurorehabilitation and human augmentation. From a systems and control perspective, NIs are closed-loop feedback systems that provide real-time decoding of neural activity and on-demand neuromodulation (nerve stimulation) of nervous circuits. The course will introduce the basic neuroscientific principles and engineering concepts required to understand and participate to this exciting field. Participants will gain insight into the generation of electrophysiological signals in the nervous system and the corresponding methods of signal processing and feature extraction. Through practical exercises participant will learn how record electrophysiological signals, and design and develop Brain-Computer interfaces (BCIs) and peripheral nervous interfaces (PNIs). PNIs are devices that directly interact with the peripheral nervous system while BCIs translate patterns of brain activity into messages for artificial devices.
Practical elements: This course is half theory and half practical. For the practical exercise, you will learn how to record electroencephalogram (EEG), electroneurogram (ENG) and electromyogram (EMG) signals and how to process and analyse the signals using Python/Matlab software.
Prerequisites/knowledge: Python/Matlab

Will I need to bring a laptop? TBC
Will I need to prepare anything in advance? TBC

Course title: Introduction to Tensorflow and Deep Learning
Provider: Dr Michael Fairbank, University of Essex
Date: 25 July 2023 (1 day)
Focus: Deep Learning
Level: Introductory level / Intermediate level
Course abstract: The course introduces Tensorflow as a programming language from scratch and shows how to use it to build simple neural networks and perform backpropagation. Students are encouraged to program along with the tutor. The basic underlying workings of TensorFlow and neural networks are taught without resorting to higher-level black box packages, so that students can gain a fundamental understanding of how deep learning works. The course also gives an introductory overview of popular deep learning models, including convolutional neural networks and recurrent neural networks.
Practical elements: The course alternates programming exercises with taught theory throughout the day. Most programming examples will be done in Jupyter notebooks and/or Google Colab.
Prerequisites/knowledge: Moderate programming experience in Python, or strong programming experience in another language.
Will I need to bring a laptop? No, delegates will have access to the University's computers. Delegate may also bring their own laptops. If bringing a laptop, you will need to be prepared with working installations of Jupyter- notebooks for running python code with the latest tensorflow python3 packages installed.
Will I need to prepare anything in advance? Optionally, undertake a python3 refresher course e.g. https://www.codecademy.com/

Course title: Tree-based Models for Machine Learning in Data Analytics
Provider: Dr Dheeraj Rathee
Date: 25 July 2023 (1 day)
Focus: Machine Learning
Level: Introductory level
Course abstract: Tree-based models e.g. Decision trees are supervised learning models used for problems involving classification and regression. These models offer a high level of flexibility that leads to an advantage towards capturing the complex non-linear relationships, however, makes these models more prone to memorizing the noise present in data. By aggregating the predictions of several decision trees that are randomly trained, ensemble methods take advantage of the flexibility feature while reducing their tendency to memorize noise. In this course, participants will learn how to work with tree-based models to solve data science problems in Python. Everything from using a single tree for regression or classification to more advanced ensemble methods will be covered. Participants start learning about basic CARTs (classification and regression trees) followed by implementation of bagged trees, Random Forests, and boosted trees using the Gradient Boosting Machine, or GBM. The course will include dedicated practical sessions for these techniques and allow the participant to create high performance tree-based models for a real-world dataset.
Practical elements:

This course is part theory and part practical. For the practical exercise, participants will learn how to implement Decision Tree, Random Forest, and Gradient Boosting algorithms to real-world datasets using Python on Google Colab platform.

Prerequisites/knowledge:
Working knowledge of Python, Jupyter Notebook and Google Colab platform.

Will I need to bring a laptop? No

Will I need to prepare anything in advance?
Participants needs to have account on Google Collab platform.

Course title: Introduction to Data Visualisation Using R

Provider: Dr Ben Skinner and Dr David Clark, University of Essex

Date: 25 July 2023 (1 day)

Focus: Advanced R Topics

Level: Introductory

Course abstract:
In the era of misinformation and fake news, producing data visualisations that are clear and interpretable to an audience is essential in engaging people with data. Whilst there are many software packages available to produce data graphics, many offer limited customisation of graphics, or are not easily reproducible. This course will explore tools to produce high-quality graphics using the R programming language, focussing on the "ggplot2" package. The "ggplot2" package allows almost endless customisation of data visualisations, has a number of excellent extension packages that add further flexibility, and, being in R, is entirely script-based and therefore highly reproducible. This course will equip attendants with the skills to produce high-quality data visualisations using the "ggplot2" package and extensions, and would be beneficial to people working in any field where data visualisation is important. The course will be suitable for those with little to intermediate prior programming experience.

Practical elements:
This course will equip attendants with the skills to produce high-quality data visualisations using the "ggplot2" package and extensions, and would be beneficial to people working in any field where data visualisation is important

Prerequisites/knowledge:
Basic knowledge of R is an advantage but not necessarily required

Will I need to bring a laptop? No

Will I need to prepare anything in advance? No

Course title: Machine Learning for Causal Inference From Observational Data

Provider: Damian Machlanski, University of Essex

Date: 26 July 2023 (1 day)

Focus: Machine Learning

Level: Intermediate

Course abstract:
This one-day workshop will:

1. Introduce the basic principles of causal modelling (potential outcomes, graphs, causal effects) while emphasising the key role of design and assumptions in obtaining robust estimates.
2. Explain the use of machine learning methods to do causal inference, and how they can be used as building blocks to obtain powerful causal estimators.
3. Show how to implement these techniques for causal analysis and interpret the results in illustrative examples.

The course covers:

- Causal modelling
- Basic machine learning techniques
- Running causal analysis on real data sets

By the end of the course participants will:

- Understand the distinction between causal effects and associations and appreciate the key role of design and possibly untestable assumptions in the estimation of causal effects
- Understand the role of training and testing models on data as well as model tuning to optimise final performance
- Be able to position machine learning within the causal tool chain.

Practical elements:
The course is roughly 30% theory and 70% practice.
Theory: basics of machine learning; fundamentals of causal modelling and analysis.
Practice: estimation and validation of causal effects using established methods, popular packages (EconML, scikit-learn) and real data sets.

Prerequisites/knowledge:
Basic knowledge of programming, ideally Python. Basics of supervised machine learning (regression and classification).

Will I need to bring a laptop, or download anything in advance?
Personal laptop is optional but recommended.
Option 1 – Google Colab
If you can access Google Colab and would like to use it throughout the practical part then there is no need to download anything in advance.
Option 2 – Local Python Environment
Alternatively, you can use a local Python environment installed on your machine. Anaconda with Python 3 is recommended in this case. We will also need EconML package for this course.

Course title: Recurrent Neural Networks with Keras

Provider: Dr Michael Fairbank, University of Essex

Date: 26 July 2023 (1 day)

Focus: Deep learning

Level: Intermediate level

Course abstract:
This course teaches a deep understanding of how recurrent neural networks work, what they are used for, and how to implement them efficiently using Keras and Tensorflow. The day culminates with unique advanced recurrent neural network examples applied to control problems. Note that natural-language processing examples will not be covered.

This course follows on nicely from the companion "Introduction to TensorFlow and Deep Learning" 1-day course; although it is also possible to take this course from scratch.

Practical elements:
The course alternates programming exercises with taught theory throughout the day. All lectures and exercises will be done using Jupyter-notebooks.

Prerequisites/knowledge:
Moderate programming experience in Python, or strong programming experience in another language.

Will I need to bring a laptop? No, delegates will have access to the University's computers. Delegate may also bring their own laptops. If bringing a laptop, you will need to be prepared with working installations of Jupyter- notebooks for running python code with the latest tensorflow python3 packages installed.

Will I need to prepare anything in advance? Optionally, undertake a python3 refresher course e.g. https://www.codecademy.com/

Course title: Bayesian Analysis in R
Provider: Dr Anna Hughes and Dr Alasdair Clarke, University of Essex
Date: 26-27 July 2023 (2 days)
Focus: R
Level: Intermediate
Course abstract: Bayesian statistics are increasingly popular in many scientific disciplines. In this course, you will learn the theoretical underpinnings of Bayesian approaches and the differences between Bayesian and frequentist statistics. You will also learn how to implement, plot, and interpret Bayesian models in R. Finally, you will learn more about some of the advanced options for statistical modelling in this framework, including multi-level modelling and generalised linear approaches.
Practical elements: This course will be half theory and half practical. We will cover the theoretical basis of Bayesian statistics, and how to implement Bayesian analyses in R, including advanced modelling techniques (e.g. multi-level models).
Prerequisites/knowledge: A basic knowledge of R (as covered in the 'Introduction to R' course).
Will I need to bring a laptop? No
Will I need to prepare anything in advance? No

Course title: The Synergy of Optimisation and Machine Learning
Provider: Dr Daniel Karapetyan, University of Nottingham
Date: 26 – 27 July 2023 (2 days)
Focus: Optimisation
Level: Introductory level / Intermediate level
Course abstract: Optimisation is a process of finding the best options given a vast number of candidates – too many to try them all directly. It is vital in decision support/making, such as finding the shortest route from Cardiff to Edinburgh, or optimising airline schedules. It is also the core of many techniques for machine learning; e.g., 'training' a neural network is actually an optimisation problem, specifically, finding connection strengths between the neurons that optimises the accuracy of its predictions. This course will teach you what optimisation is, where it occurs, including its use in machine learning), and how we can address optimisation problems. It will also show how machine learning improves modern optimisation methods. In the first part of the course, we will discuss the meaning of optimisation and its relevance to decision support/making and machine learning, computational complexity, and the implications for the difficulty of optimisation, simple heuristics and metaheuristics. The second part of the course will provide an overview of more advanced topics of optimisation and the strong connections between optimisation and data science. We will study more examples of optimisation problems and solution techniques used in practice. We will also see how machine learning is increasingly used within optimisation methods to automate the process of building optimisation algorithms. The lectures will be interleaved with practical exercises on both days.

Practical elements:
To give hands-on experience, the lectures will be accompanied by lab exercises including running and modifying existing programs (mostly in Python) to obtain practical experience of modelling and applying basic optimisation techniques, as well as understand the effects of using more advanced approaches.

Prerequisites/knowledge:
A background in an analytical discipline such as computer science, mathematics, or engineering. Basic knowledge of some machine learning technique(s) will be useful. To participate in lab exercises, one will need programming skills sufficient to modify simple code, though solutions will be provided and demonstrated to the class.

Will I need to bring a laptop? Yes – for working on the exercises. If you do not have a laptop, you will be able to group up with other students who have laptops or observe the demonstrations on the big screen.

Will I need to prepare anything in advance? If you bring your laptop, please make sure that it has Python 3.7+ and an IDE of your choice, for example PyCharm.

Course title: Learning Under Different Training and Testing Distribution

Provider: Mohsin Ali, University of Essex

Date: 27 July 2023 (1 day)

Focus: Deep learning

Level: Intermediate level

Course abstract:
Systems based on machine learning methods often suffer a major challenge when applied to the real-world datasets. The conditions under which the system was developed will differ from those in which we use the system. Few sophisticated examples could be email spam filtering, stock prediction, health diagnostic, and brain-computer interface (BCI) systems, that took a few years to develop. Will this system be usable, or will it need to be adapted because the distribution has changed since the system was first built? Apparently, any form of real-world data analysis is cursed with such problems, which arise for reasons varying from the sample selection bias or operating in non-stationary environments. This tutorial will focus on the issues of dataset shifts (e.g. covariate shift, prior-probability shift, and concept shift) and will cover transfer learning for managing to learn a satisfactory model.

Practical elements:
This course is half theory and half practical. For the practical exercise, you will learn how to learn under different training and testing distributions.

Prerequisites/knowledge:
Working knowledge of Python and GoogleColab.

Will I need to bring a laptop? No

Will I need to prepare anything in advance?
Participants needs to have account on Google Colab platform.

Course title: Intro to Startups for Data Scientists, Analysts and ML Engineers

Provider: Dr Tom Vodopivec, Silicon Gardens Fund

Date: 27 July 2023 (1 day)

Focus: Cross theme

Level: Introductory

Course abstract:
An introductory course designed for technical professionals and students (data scientists, analysts, ML engineers, software developers) who are interested in understanding the world of startups. The course will cover the fundamentals of startups and how they function, the career opportunities available, and how to start your own startup. Participants will learn about the role

of venture capital, pitch decks, business model canvases, and the concept of control and cap tables.

Practical elements:
Through practical exercises, attendees will explore career paths that align with their mindset and life goals. how to join a startup, how to start their own startup. This will include case studies, discussions, hands-on exercises (drafting a business based on ideas), and a Q&A session with an experienced startup advisor and investor.

Prerequisites/knowledge: none, but some pre-preparation is encouraged (see below).

Will I need to bring a laptop? Ideally yes, because during the practical exercises you will be extensively searching the web for additional info, otherwise you could also do this on your smartphone but you will likely be much less efficient.

Will I need to prepare anything in advance?
Register an account for FigJam/Figma.

Non-obligatory preparation prior the course:

1. Do a brief mental exercise: fast forward 5 years, if someone asks you what have you achieved in this time, what would be your ideal answer? And, if in 5 years someone asks you what are you doing today, what is your job and responsibilities, what would be your ideal answer?
2. Find a startup or tech company that you admire and research a bit about it: what is the product, what are the customers, how big it is, how long it took them to build this business, who are the founders, the investors (shareholders), etc. You can also pick a startup that you might be considering joining or a startup an acquaintance or friend of yours is starting up. Put the notes into a document of a couple of pages.
3. If you have an idea of your own or have been working on a passion project, do a similar exercise as in the previous point - write a page (or a few of them) on what kind of startup this would be.

Course title: Introduction to Federated Machine Learning

Provider: Dr Salman Toor, Uppsala University and Dr Mays AL-Naday, University of Essex

Date: 28 July 2023 (1 day)

Focus: Cross theme

Level: Introductory/intermediate

Course abstract:
Federated machine learning has opened new avenues for privacy-preserving data analysis. Instead of pooling data in a central location, different data owners keep data local and training is decentralized where only model parameters are exchanged. This course will explore the range of training and networking challenges in relation to model exchange among client and server nodes. We will also cover the recent developments in the field and some well know solutions to enhance privacy and security in federated machine learning. The second half of the day will entail hands-on experiments using state-of-the-art open-source European federated learning platform, FEDn, which will be hosted over the Network Convergence Laboratory (NCL) in UEssex. The FEDn is a lightweight, scalable, and efficient framework that is vendor-agnostic both in terms of the underlying machine learning libraries and also the distributed infrastructure technologies. The foundation of the FEDn architecture is based on the map-reduce paradigm, a well-known scalable distributed systems design. Based on the proposed theoretical and practical sessions, the participates will have a good understanding of the topic and it will provide the foundation required to further explore the field.

Practical elements:
The course will be half theory and practical. In the first half, there will be two lectures (09:30 – 10:30 and 11:00 – 12:00). After the lectures, there will be a short discussion session. In the second

half, student will work on the FEDn framework. The hands-on part will be for 3 hours (13:30 – 17:00).

Prerequisites/knowledge:
Interest in data science and machine learning and distributed systems. Students also preferable if they have some knowledge of linux commandline tools.

Will I need to bring a laptop? Yes

Will I need to prepare anything in advance?
They need to have a working docker environment on their laptops and vpn application that supports openvpn (e.g. tunnelblick for mac and OpenVPN for windows) to be able to connect to NCL.

Course title: Empowerment – an Information Theoretic Intrinsic Motivation

Provider: Dr Christoph Salge, Karen Archer

Date: 28 July 2023 (1 day)

Focus: Cross theme

Level: Introductory to intermediate

Course abstract:
Empowerment measures how much an agent is in control of the world it itself perceives. This information-theoretic intrinsic motivation measure was originally conceived to understand the development of more complex behaviours via simple gradients, but has since spread to many other domains. In this course we will first cover the core idea of Empowerment, introducing basic concepts of information theory, intrinsic motivation and the perception action loop. After a short practical exercise we will then take a look at some neural network approximation, and the application of empowerment to fields such as deep learning, robot control and multi-agent reinforcement learning.

Practical elements:
The course will be split in four, roughly equally long parts:

1. An introductory presentation on empowerment and core related concepts, such as information theory, perception action loop, etc. This will be in detail, and partially interactive, to enable participants to implement the concepts.
2. A coding exercise, which will extend upon existing code, like going to be provided via something like Jupyter Notebook.
3. A more abstract lecture outlining the development of empowerment heuristics and a few modern example applications.
4. A discussion regarding questions, and possible future applications of empowerment.

Prerequisites/knowledge:
Background in probability theory, information theory, or control theory is helpful, but not necessary. An interest to look at the bridge between cognitive and social phenomena and quantitative methods to generate them is helpful. This concept has originated from the Artificial Life field.

Will I need to bring a laptop? Yes

Will I need to prepare anything in advance? No

Course title: Best Practice Analytics

Provider: Dr Detlef Nauck, BT

Date: 28th July 2023 (1 day)

Focus: Cross theme

Level: Intermediate level

This course will look at methods and tools that can help us create high-quality analytics and reproducible results. We will also look at how to move from a single analyst, spreadsheet driven

approach to collaborative analytics that follows a best practice governance model. Adopting practices from test driven software development, we will look at how to establish an analytics process based on documentation, versioning, testing, peer review, collaboration and risk evaluation. We will use examples in R, Shiny, Python and Jupyter Notebooks to illustrate the ideas taught in the course.
The aim is to give you an understanding of the challenges you will face when running your own real-world data analytics project and introduce you to a number of principles you can follow to achieve high-quality reproducible results.

You will learn

- how to use ideas from test driven development to control the quality of your data analysis including checks for bias and fairness
- how to write assertions on data to make sure it is correct before you use it for analysis and machine learning
- how to set up machine learning experiments and evaluate the performance and reliability of your results
- - how to interpret the outcomes of analytics and present them to business stakeholders.

Practical elements:
The course will be half theory and half practical exercises. In the exercises you will learn how to build a data analysis / machine learning pipeline visualise the results and create an automatic report. You can pick between an R and a Python version of the exercises.

Prerequisites/knowledge:
A basic knowledge or R or Python is useful, but not required. Interest in machine learning.

Will I need to bring a laptop? Yes

Will I need to prepare anything in advance?
Get a free Google Account so you can use https://colab.research.google.com/
Alternatively, pick any Jupyter Notebook server of your choice.
You can also install R and/or Python on your laptop and work locally.

Course title: Introduction to Network Science

Provider: Professor Philip Leifeld

Date: 28 July (1 day)

Focus: Cross theme

Level: Introductory

Course abstract:
This one-day introductory course on network science will give a broad overview of the different concepts and methods commonly applied in social network analysis. We will first consider different kinds of network data and their representation and discuss the basics of network visualisation, including a hands-on example using the free software visone. We will also discuss different kinds of applications and usage scenarios of network science in business and social contexts. The second part of the course will introduce exploratory and descriptive methods for the analysis of networks, at three levels of granularity: at the node level, the subgroup level, and the network level. The third part of the course will introduce inferential or statistical network analysis, including the basic ideas behind a range of models like the exponential random graph model and its various extensions, latent space models, the quadratic assignment procedure, and related techniques. We will cover the implementation of these methods in a very cursory way using R, but the focus is on the methods, not their implementation. Overall, this course is an introductory-level teaser for interested academics, practitioners, and data scientists who would like to explore what they can possibly do with their relational data in the way of exploration and prediction.

Practical elements: This course gives a broad overview of different methods and models, often using examples. Methods are thus comprising about two thirds of the course. We will also consider software implementations, which is about one third of the course. This third is split into one half in visone (no prior skills necessary) and the statnet and igraph packages in R (no prior skills necessary).
Prerequisites/knowledge: R skills desirable but not necessary. Statistical skills at the level of logistic regression desirable but not necessary.
Will I need to bring a laptop? Yes.
Will I need to prepare anything in advance? Installation of the (free) software as outlined below under software access.

Course title: GIS Systems in R
Provider: Dr Ben Skinner and Dr David Clark, University of Essex
Date: 28 July 2023 (1 day)
Focus: Advanced R Topics
Level: Intermediate
Course abstract: Geographic information systems (GIS) software form a powerful tool in the analysis of many types of spatial data, from understanding political trends in different areas, mapping the spread of infectious diseases, or understanding the impacts of climate change across the globe. Commonly used GIS tools offer great power, but can be incredibly expensive for individual users, or offer limited reproducibility of analyses. In recent years, the "landscape" of GIS packages available in the R programming language has enabled R to become a powerful and richly-functional tool in the world of GIS anaysis. R, and all of it's packages, are freely available, and entirely script-based, allowing users to quickly and easily reproduce their analyses. This course will focus on the "sf" package, and will explore the merits and functionality of working with "simple features" based objects in geographical analyses.
Practical elements: This one-day course will familiarise users with the array of GIS packages available in R, and enable users to carry out basic GIS operations on a variety of different geographical data formats.
Prerequisites/knowledge: Some prior experience with geographical data and GIS software would be beneficial, but is not essential, whilst this course is recommended for users with an intermediate level of programming experience (or who have attended the "Introduction to R" and "Data visualisation in R" courses).
Will I need to bring a laptop? No
Will I need to prepare anything in advance? No

Course Provider Biographies

Mohsin Ali is a PhD student in CSEE Department at the University of Essex. His PhD is related to Explainable AI in Medical Imaging. His research experience field of expertise is Computer vision and Image processing.

Karen Archer is a PhD candidate at the University of Hertfordshire where she works on the geometry of cognitive decision making, together with Prof. Daniel Polani – one of the inventors of Empowerment.

Dr Dave Clark is a current Research Fellow within the Institute for Analytics and Data Science at the University of Essex and achieved a PhD in Microbiology at the University of Essex in 2017. His research involves combining bioinformatic-, statistical- and geographical-analyses to answer novel questions on global microbial community ecology. Dave has extensive experience teaching programming skills in R at both undergraduate and postgraduate levels, in addition to the multidisciplinary programming skills required by his research programme.

Dr Alasdair Clarke is a senior lecturer in the Department of Psychology at the University of Essex. He has a background in mathematics and computer science and his research interests involve computational modelling of visual search and decision making.

Dr Anirban Chowdhury is a Lecturer in Neural Engineering and Robotics, School of ComputerScience and Electronic Engineering (CSEE), University if Essex (UoE) and a member of the Brain-computer interface and Neural Engineering (BCI-NE) Group, and Robotics Group at University of Essex. Prior to the joining at UoE Anirban was a postdoctoral Research Associate at the Northern Ireland Functional Brain Mapping Facility at Ulster University, Northern Ireland, UK. His current research interests are in the areas of robotic rehabilitation, brain-computer interfaces, assistive technologies, human-robot co-operation, and autonomous mobile robotics. He has published several journal papers in the top journals of the field, including transactions and journals of the IEEE and Elsevier.

Dr Ian Daly is a Senior Lecturer in the School of Computer Science and Electronic Engineering at the University of Essex. He received an M.Eng. degree in Computer science in 2006 and a Ph.D. degree in Cybernetics in 2011 from the University of Reading. Between May 2011 and 2013 he worked as a post-doctoral researcher in the Laboratory of Brain-Computer Interfaces, Graz University of Technology, Graz, Austria. There he researched Brain-computer interfaces (BCIs) for individuals with stroke and cerebral palsy. He then worked for three years in the University of Reading developing BCIs for music therapy.

Dr Sebastian Halder is a Senior Lecturer in Brain-Computing Interfacing at the University of Essex. His research interests include auditory BCIs for communication, the neural mechanisms of learning with BCIs, the neural signature of pain and disorders of consciousness.

Dr Audrey Guinchard is a senior lecturer in the School of Law. Her current research is at the intersection of cybercrime, data protection and cybersecurity.

Dr Mays AL-Naday is an Assistant Professor at the School of Computer Science and Electronic Engineering, University of Essex (UEssex). Her research expertise include service-oriented networking, Fog computing networks and machine learning application in cybersecurity. AL-Naday is currently the managing academic of the Network Convergence Laboratory (NCL), she has been the co-PI on the H2020 project SerIoT and a senior research officer on the H2020 project POINT. AL-Naday's past engagements relevant to this project include the organization of high quality

workshops along with Sigcomm and IFIP NETWORKING conferences and development team leader within the POINT project.

Dr Michael Fairbank is a senior lecturer in computer science, at the University of Essex. He is an active machine-learning researcher, with publications in reinforcement learning, deep learning and neural networks. In his previous careers he worked as a computer consultant and as a mathematics teacher. He has a passion for all things related to computing, mathematics and AI.

Dr Anna Hughes is a lecturer in the Department of Psychology at the University of Essex. Her research focuses on visual perception and decision making, asking how humans and other animals search the world. She has extensive experience with R, and teaches statistics and programming to Masters students.

Dr Daniel Karapetyan is an assistant professor at the University of Nottingham. He received his PhD from Royal Holloway, University of London in 2010 and then worked at the Simon Fraser University (Canada), University of Nottingham, and finally the Institute for Analytics and Data Science (IADS) at the University of Essex before taking his current role. Daniel's main expertise is algorithm design for decision support. He mainly focuses on artificial intelligence, optimisation techniques, machine learning and automated algorithm design. He participated in many applied projects with large businesses, working on ferry scheduling, airport and airline operations, satellite mission planning and others.

Dr Sefki Kolozali is a Lecturer in Embedded and Intelligent Systems at the University of Essex and part of the Robotics and Embedded Systems Research Group. He is interested in the fields of Internet of Things and Big Data Analytics.

Prof Philip Leifeld is a Professor of Comparative Politics in the Department of Government and Deputy Director of the Centre for Social and Economic Network Analysis (SENA) at the University of Essex. Before joining Essex in 2019, he was a Professor of Research Methods at the University of Glasgow in the School of Social and Political Sciences. He holds a PhD in Politics and Public Administration from the University of Konstanz and completed research posts at the Max Planck Institute for Research on Collective Goods, the Max Planck Institute for Demographic Research, the Swiss Federal Institute of Aquatic Science and Technology (Eawag), and the University of Bern. Professor Leifeld is the author of several software packages, including the Java software Discourse Network Analyzer and the R packages rDNA, texreg, and btergm. His research was published in leading journals in several fields, spanning the social sciences and STEM fields. Professor Leifeld has contributed to the development of network analysis methodology and its application to the study of politics and other fields. His current research is funded by the Economic and Social Research Council and Innovate UK.

Damian Machlanski is a computer science Ph.D. student, jointly under CSEE and MiSoC (ISER), University of Essex, working at the intersection of machine learning and causality with Dr Spyros Samothrakis and Prof. Paul Clarke, with a particular emphasis on the methods of causal effect estimation, causal discovery, treatment recommendation, policy learning as well as model tuning and validation -- all using observational data. Prior to his current position, he worked as a software developer for several years and obtained his MSc in Artificial Intelligence at the University of Essex.

Dr Detlef Nauck is Distinguished Engineer and Head of AI & Data Science Research at BT. Detlef has over 30 years of experience in AI and Machine Learning and leads a programme spanning the work of 30 international researchers who develop capabilities underpinning modern AI systems. A key part of the work is to establish best practices in Data Science and Machine Learning for conducting data analytics professionally and responsibly leading to new ways of analysing data for achieving better insights. Detlef is a computer scientist by training and holds a PhD and a Postdoctoral Degree

(Habilitation) in Machine Learning and Data Analytics. He is a Visiting Professor at Bournemouth University and a Private Docent at the Otto-von-Guericke University of Magdeburg, Germany. He has published 3 books, over 120 papers, and holds over 20 AI patents.

Dr Dheeraj Rathee is currently holding the position of Chief Technology and Analytics Officer at one of the UK's fast-growing HealthTech start-ups - Provide Digital. Dheeraj has been endorsed as Global Exceptional Talent in the field of AI and Data Analytics by Tech Nation in 2021. He received his PhD degree in Computer Science from Ulster University, UK in 2018. He received his bachelor's degree in Biomedical Engineering and master's degree in Electronics and Communication Engineering in 2007 and 2011. Dr Rathee has major experience in developing and deploying AI and Data-driven solutions involving innovative technologies such as Natural Language Processing, Time-series Analytics, Supervised and Unsupervised ML, Predictive Analytics, and cloud computing in application domains of healthcare, finance, and social science. He has published more than 20 research papers in peer-reviewed journals and conference proceedings.

Dr Christoph Salge is a Reader in AI in Games at the University of Hertfordshire. He has previously completed a Marie-Curie Global Fellowship at NYU, working on AI and Games. Previously he has worked on two EU projects that drove the development of Empowerment, and applied to the control of rehabilitation robots and social entrainment.

Prof Reinhold Scherer is a Professor in Brain-Computer Interfaces (BCIs) and Neural Engineering (NE) in CSEE and co-director of the Essex BCI-NE laboratory. His primary research interests are in the areas of online brain-machine co-adaptation, statistical and adaptive signal processing, mobile brain and body imaging, and rehabilitation.

Dr Ben Skinner obtained a PhD in Genetics from the University of Kent in 2009, and then performed postdoctoral research at the University of Cambridge on structural and evolutionary genomics - how genomes and the chromosomes they contain change and rearrange over time. In 2019 he joined the University of Essex as a Lecturer in the School of Life Sciences, where he conducts research into the organisation of chromatin in asymmetric cell nuclei such as sperm which can give information on mechanisms of chromatin packaging.

Dr Salman Toor is Associate Professor in Scientific Computing at Uppsala University. He is an expert in the field of distributed computing infrastructures and applied machine learning. Toor is the co-chair of the Integrative Scalable Computing Lab (ISCL) at Uppsala University and co-founder and CTO at Scaleout Systems AB. He has supervised and reviewed more than forty master thesis projects and is currently the principal supervisor of a Ph.D student. Toor is also one of the lead architects of the FEDn framework, designed and developed for scalable federated machine learning. He was project leader for the project, "A Platform for Privacy-preserving Machine Learning Using the Ethereum Blockchain and Smart Contracts" and currently, he is leading Scaleout's activities in the SESAR funded AICHAIN project. Toor is a co-PI of the Gigacow project and a named researcher in the HASTE and SustAinimal projects. Together with extensive research experience, Toor has over 15 years of teaching experience in different international institutions. He has designed and developed both masters and PhD level courses in the field of distributed computing infrastructures. From the platform of Scaleout Systems, Toor has organized a number of national and international workshops and tutorials.

Dr. Tom Vodopivec, a data scientist by background, is a venture capital principal, startup advisor, and mentor in entrepreneurship programs across Europe. He co-founded and directed the department for analytics and data science of the worldwide-leading advertising technology startup Celtra, he founded the startup studio RV2 ventures, and helped more than 40 companies develop their business through data-driven decisioning. Today within Silicon Gardens he is managing and

setting up venture funds for early-stage startups. He holds a PhD from artificial intelligence - namely from reinforcement learning, search algorithms, and neural networks applied to decision problems.