SIMA-SIM-BAM
SCHOOL OF RESEARCH METHODS
Summer Edition: *An introduction to Python, Network Analysis and Machine Learning*

Certosa di Pontignano (Siena, Italy)
and
Department of Business and Law (University of Siena)
17th-20th July 2024

Programme

Welcome, Registration and Check-in

17th July 2023, 14:00-14:30
Department of Business and Law, University of Siena

Giovanni Battista Dagnino – University of Rome LUMSA
Rossella Chiara Gambetti – Università Cattolica del Sacro Cuore
Elena Casprini – University of Siena

14:00 Arrival at Siena, gathering and registration of participants at the Department of Business and Law of the University of Siena, Piazza San Francesco 7/8, Siena

14:30 Leave for Certosa di Pontignano by bus

15:00 Arrival at **Certosa di Pontignano** and room check-in

16:00 All physically ready and mentally prepared for the opening seminar!

Complementing Python with ChatGPT

17th July 2023, 16:00-19:00

Certosa di Pontignano

Gian Piero Cervellera is lecturer on contract of Applied Computer Science at the University of Siena. An expert in Artificial Intelligence, he is a consultant in AI at Aubay, where he coordinates the Data Scientist group. Co-owner of FootballIntelligence—a company aimed at transforming all the statistics generated in the world of football and beyond into useful information for coaches and their staffs—he is the head of the Data Scientist group.

Content

This lesson will introduce the use of Python for programming, leveraging ChatGPT to facilitate code writing. Even those approaching Python for the first time will be able to create programs. We will use Google Colab, a platform integrated within Google Drive, therefore a Google account is required. Access to ChatGPT (free version) is also required. During the three-hour course, we will explore the use of Artificial Intelligence models available on Huggingface (https://huggingface.co/), for which a free account on this platform is necessary. The goal is to understand how to employ these models in Google Colab through ChatGPT, enabling students to develop comprehensive AI programs for various applications, from computer vision to natural language processing, suitable for any work context.

Teaching materials

Presentations and references will be provided by the instructor.

Digital Network Analysis
Theory
18th July 2023, 9:30-12:30
Certosa di Pontignano

Alessandro Caliandro is an Associate Professor in Sociology of Culture and Communication at the Department of Political and Social Sciences at the University of Pavia. Previously he worked as Assistant Professor in Digital Marketing at the Bath School of Management, University of Bath. His current research focuses on digital methods, digital consumer culture, platformisation of culture, surveillance capitalism, ageing and digital media. He is Scientific Coordinator of V-DATA (https://vdataresearch.com/), which focuses on the critical study of surveillance capitalism in Italy, and Research Stream Leader of Algofeed (Feedback culture: assessing the effects of algorithmic recommendations on platformized consumption). His works have been published in international journals such as Journal of Consumer Research, Marketing Theory, Journal of Marketing Management, New Media & Society, Information, Communication & Society, and Social Media+ Society. He is author of 'Qualitative Research in Digital Environments: A Research Toolkit' (Routledge, 2017), and 'The Platformization of Consumer Culture: A Digital Methods Guide' (Amsterdam University Press, 2024).

Content
This course aim at illustrating students how social network analysis is employed within the paradigm of digital methods. Rather than focusing on social hierarchies and structures, digital methods take advantage of network analysis to map online communities, discourses, and the digital structure of online environments. This approach amounts to be particularly useful for consumer and marketing scholars interested in exploring online cultural processes.

This module is about Theory:
- Introducing the digital methods paradigm
- Digital network analysis: basic vocabulary and metrics
- Case studies: answering relevant research question with network analysis

Digital Network Analysis
Practice
18th July 2023, 15:30-18:30
Certosa di Pontignano

Ilir Rama is a post-doctoral researcher at the Department of Social and Political Sciences at the University of Milan, where he works on the ALGOFEED project (Feedback culture: assessing the effects of algorithmic recommendations on platformized consumption). His work focuses on the social and cultural impact of social media, algorithms, and generative artificial intelligence, at the intersection of computational and qualitative methods. He previously collaborated with Tracking Exposed (now AI Forensics) auditing a variety of algorithmic systems.

Content

This second module is about Practice:
- Introducing Gephi and NetworkX
- From unstructured data to network datasets
- Visualizing social formations through Gephi

Machine learning to investigate complex systems
19th July 2023, 9:30-12:30
Certosa di Pontignano

Michele Tumminello is a Professor of Applied Mathematics at the University of Palermo. His research spans complexity, network science, bioinformatics, and textual analysis. He is currently dedicated to developing machine learning and network techniques tailored for big data applications in criminology, with a focus on organized-crime, and frauds in the insurance sector. He has authored over 50 articles published in international journals, including PNAS, BMC Bioinformatics, Journal of Information Science, Research Policy, Social Networks, Quantitative Finance, Journal of Risk and Insurance, and Journal of Economic Behavior and Organization.

Content
Machine learning (ML) systems learn and make decisions based on patterns and insights extracted from data. ML encompasses a variety of techniques and approaches, from logistic regression (supervised learning) to clustering (unsupervised learning). As a technology that adapts well to a broad range of data types, ML is increasingly employed across diverse sectors such as finance, management, economics, healthcare, and the social sciences, revolutionizing the paradigm for analyzing and modeling complex systems—from understanding "why" something happens to predicting "how" and "when" it will occur. In this lecture, I will adopt a bottom-up approach to explain the fundamentals of machine learning—demonstrating how the narrative sections of corporate reports can inform an ML system to predict firm bankruptcy.

Teaching materials
Presentations and references will be provided by the instructor.
Preliminary (useful but not mandatory) readings:
- Vishal Maini and Samer Sabri (2017) Machine Learning for Humans (free for download at https://medium.com/machine-learning-for-humans/why-machine-learning-matters-6164faf1df12)

Implementation of simulated annealing for community detection through modularity optimization in complex networks.

19th July 2023, 15:00-18:00

Certosa di Pontignano

Michele Tumminello is a Professor of Applied Mathematics at the University of Palermo. His research spans complexity, network science, bioinformatics, and textual analysis. He is currently dedicated to developing machine learning and network techniques tailored for big data applications in criminology, with a focus on organized-crime, and frauds in the insurance sector. He has authored over 50 articles published in international journals, including PNAS, BMC Bioinformatics, Journal of Information Science, Research Policy, Social Networks, Quantitative Finance, Journal of Risk and Insurance, and Journal of Economic Behavior and Organization.

Content

In this hands-on session, participants will be guided through the implementation of a machine learning system aimed at detecting communities in a complex network through modularity optimization. Optimization involves maximizing or minimizing a utility or a cost function, respectively, and represents the core element/process in model training–one of the fundamental aspects of machine learning. Currently, modularity optimization is still widely used to detect communities in complex networks, though other methods revealed to be more effective in some cases. The optimization will be carried out by using an heuristic algorithm– the simulated annealing (SA). Inspired by the process of annealing in metallurgy, where materials are heated and then slowly cooled to increase their ductility and reduce defects, SA is an effective optimization algorithm in presence of a large and rugged solution space. The programming language used will be R; however, only a basic understanding of programming concepts is required.

Teaching materials

Presentations, references, and codes (in R and Python) will be provided by the instructor. Preliminary (useful but not mandatory) readings:

- Alex Douglas, Deon Roos, Francesca Mancini, Ana Couto & David Lusseau (2024) *An Introduction to R* (free for download at https://alexd106.github.io/intro2R/)
- Mark E.J. Newman (2006) *Modularity and community structure in networks*. Proceedings of the national academy of sciences, 103(23), 8577-8582. (free for download at https://www.pnas.org/doi/abs/10.1073/pnas.0601602103)
- Gendreau, M., & Potvin, J. Y., (2010). *Handbook of metaheuristics*. Vol. 272, New York: Springer. (free for download at https://hughchristensen.com/papers/academic_papers/Handbook%20of%20Metaheuristics%202nd%20Edition.pdf)

Final day
20th July 2023, 9:00
Certosa di Pontignano

9:00 Room check-out, gathering and leave by bus for Siena
9:30 Arrival in Siena, Department of Business and Law, Piazza San Francesco 7/8

Computational text analysis in organization and management theory
20th July 2023, 10:00-13:00
Department of Business and Law, University of Siena
Piazza San Francesco 7/8, Siena

Simone Santoni is Senior Lecturer (associate professor) in Strategy at Bayes Business School (London, UK). His research combines computational methods and sociological arguments to appreciate the functioning of markets—especially markets for cultural products—and open, collaborative organizations. Throughout his work, he emphasizes the interrelationship of individuals, groups and social structures and tries to detail the causal mechanisms through which organizational and social facts are brought about.

Content.
The session deals with computational text analysis' foundations and scope of application to organization and management theory. By reviewing concrete examples, the first part of the session will clarify the gamut of computational text analysis tools and individual tools' distinctive features and strengths. The second part will focus on Topic Modeling as a popular and versatile computational text analysis model. Specifically, the attention will revolve around Topic Modeling's use cases, training, and validity assessment.

Teaching materials
Presentations and references will be provided by the instructor.