

Course proposal

Explainability for symbiotic AI

Teacher(s)

Gianluca Zaza

Course description (min 150, max 300 words)

While effective, most of the current prominent AI methods consist of "opaque" models that are inherently complex to interpret and explain. The intricacies of deep neural networks, complex statistical models, and reinforcement learning algorithms often result in decision-making processes that appear as "black boxes" to the end user. This opaqueness poses significant challenges in scenarios where AI systems impact our lives, from healthcare diagnosis and autonomous vehicles to financial risk assessment and recommendation engines. Designing and developing practical explanation components for data-driven learning models is imperative to cope with this need for interpretability. These explanation components are a crucial bridge between the AI system's decision-making processes and the human users who rely on these systems. Such components enhance transparency, trust, and user acceptance in AI applications across diverse domains. These explanation components go beyond mere model visualization or summary statistics. They delve into the inner workings of AI algorithms, breaking down complex decision paths into understandable and human-friendly formats. The goal is to make AI systems comprehensible and empower users with insights into how and why specific decisions are made. These components provide meaningful explanations and assist users in decision-making processes and knowledge-intensive tasks. They enable healthcare professionals to trust diagnostic recommendations, help drivers understand why an autonomous vehicle made a particular manoeuvre, guide traders in comprehending trading decisions, and offer tailored product recommendations while justifying their relevance. In the ever-evolving landscape of artificial intelligence, achieving interpretability is an ongoing challenge. The course will cover various topics, including techniques for both ante and post hoc explainability and explainability in hybrid models.

Course period

November-December 2023

SSD

INF/01

Course References (optional)

- [1] Arrieta, Alejandro Barredo et al. "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI." *Inf. Fusion* 58 (2019): 82-115.
- [2] Molnar, C. (2022). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable* (2nd ed.). christophm.github.io/interpretable-ml-book/
- [3] Denis Rothman, "Hands-On Explainable AI (XAI) with Python: Interpret, visualize, explain, and integrate reliable AI for fair, secure, and trustworthy AI apps", Packt Publishing Ltd, 2020.

Credits and Hours

3 CFU, 2 of lecture (8 Hours) and 1 of practice (15 hours), for a total of 31 hours.

Exam Modality

Two alternatives are available to the student to pass this exam (Teacher(s) may choose other modalities):

- 1) Paper presentation. Students present the content of 2 papers suggested by the teacher. No groups are allowed.
- 2) Project. Students implement and experimentally validate an algorithm or its variation from a paper suggested by the teacher. Projects can be done in groups of 1-3 students, depending on the algorithm.

Teacher(s) CV

Attach or link a max 3 pages CV for each teacher proposing the course.

PERSONAL INFORMATION

Family name, First name: Zaza, Gianluca

Researcher unique identifier(s) (such as ORCID, Research ID, etc.

...): <https://orcid.org/0000-0003-3272-9739>

Date of birth: 04/09/1992

Nationality: Italian

URL for web site: <https://sites.google.com/site/cilabuniba/people/gianluca-zaza>

• EDUCATION

- 2022 PhD in Computer Science, University of Bari Aldo Moro, Bari-Italy.
Name of PhD Supervisor: prof. Giovanna Castellano
- 2017 Master's degree in Computer Science, University of Bari Aldo Moro, Bari-Italy.
Name of Faculty/ Department, Name of University/ Institution, Country
- 2015 Bachelor in Computer Science and Digital Communication, University of Bari Aldo Moro, Bari-Italy.

• CURRENT POSITION(S)

- 2023-2026 Assistant Professor (non-tenure track), on the funds of the project FAIR - Future AI Research (PE00000013), Spoke 6 - Symbiotic AI (CUP H97G22000210007), under the NRRP MUR program funded by the NextGenerationEU, Computer Science Department University of Bari Aldo Moro, Bari, Italy,

• PREVIOUS POSITION(S)

- 2022-2023 Contract researcher with a research grant with title: "Progettazione di modelli per lo smaltimento dell'arretrato civile con tecniche di Intelligenza Artificiale", on the funds of the project "Start UPP - Modelli, sistemi e competenze per l'implementazione dell'ufficio per il processo", Computer Sciences Department, University of Bari Aldo Moro, Bari-Italy.

• FELLOWSHIPS AND AWARDS

- 2018-2022 PhD scholarship, Computer Sciences Department, University of Bari Aldo Moro, Bari, Italy.
- 2021 Best Paper Award received at the Doctoral Consortium Symposium on Computational Intelligence, presented in the Institute of Engineering and Technology, a constituent college of Dr APJ Abdul Kalam Technical University Lucknow, India.
- 2019 Vision and Imaging Technology Award: Call for Demos, presented during the 9th International Conference on Imaging for Crime Detection and Prevention, University of Westminster, London.

- **TEACHING ACTIVITIES**

- 2021-2024 Subject matter expert, teaching Intelligent Systems, Bachelor Degree Course in Computer Science for Digital Communication, Department of Computer Science, University of Bari Aldo Moro, Taranto Campus.
- 2018-2020 Teaching support activities, as a teaching load for the PhD course, for the courses of Computer Science Laboratory, Programming Languages, in Bachelor's degree course in Computer Science and Digital Communication and Bachelor's Degree in Computer Science and Technologies for Software Production, University of Bari Aldo Moro.

- **ORGANISATION OF SCIENTIFIC MEETINGS (if applicable)**

- 2023 Organizer of the Special Session with title: "New frontiers of Computational Intelligence for Pervasive Healthcare Systems", during the 13th Conference of the European Society for Fuzzy Logic and Technology EUSFLAT, Palma (Spain).
- 2022 Organizer of the Special Session with title: "New frontiers of Intelligent Pervasive Healthcare Systems" during the International Conference on Systems, Man, and Cybernetics (SMC), Prague (Czech Republic).

- **INSTITUTIONAL RESPONSIBILITIES**

- 2019-2023 Co-supervisor for six bachelor's and master's theses, with the main topics being Artificial Intelligence, Mobile Apps, Computer Vision, and Healthcare.

- **REVIEWING ACTIVITIES**

- 2023 Associate Editors - Journal of Intelligent & Fuzzy Systems (IOS press).
- 2022 Guest co-editor - Special Issue "Computational Intelligence in Healthcare", Bioengineering (MDPI) Guest co-editor della Special Issue "Computational Intelligence in Healthcare", Bioengineering (MDPI).
- 2023-2021 Technical Program Committee member for numerous international conferences such as: AI4eHealth 2023, ISCC, CVCI 2023, ICTS4Health.

- **MEMBERSHIPS OF SCIENTIFIC SOCIETIES**

2019-2023 Member of INDAM-GNCS, Gruppo Nazionale per il Calcolo Scientifico.

- **MAJOR COLLABORATIONS**

Prof. Uzay Kaymak, Development of interpretable intelligent systems based on neuro-fuzzy models in the healthcare sector. Eindhoven University of Technology, Netherlands.

Teacher(s) Main Publications

List 10 main publications in the last 15 years for each teacher.

[1] Casalino, G., Castellano, G., **Zaza, G.** Evaluating the robustness of a contact-less mHealth solution for personal and remote monitoring of blood oxygen saturation. *Journal of Ambient Intelligence and Humanized Computing*, 2023, 14(7), pp. 8871–8880.

[2] Casalino, G., Castellano, G., **Zaza, G.** Interpretable Neuro-Fuzzy Models for Stress Prediction. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2023, 14069 LNCS, pp. 630–641.

[3] Casalino, G., Castellano, G., **Zaza, G.** Explainable Fuzzy Models for Learning Analytics. *Lecture Notes in Networks and Systems*, 2023, 716 LNNS, pp. 394–403.

[4] Casalino, G., Castellano, G., Nisio, A., Pasquadibisceglie, V., **Zaza, G.** A mobile app for contactless measurement of vital signs through remote Photoplethysmography. *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, 2022, 2022-October, pp. 2675–2680.

[5] Casalino, G., Castellano, G., **Zaza, G.** Neuro-Fuzzy Systems for Learning Analytics *Lecture Notes in Networks and Systems*, 2022, 418 LNNS, pp. 1341–1350.

[6] Casalino, G., Castellano, G., **Zaza, G.** On the use of FIS inside a Telehealth system for cardiovascular risk monitoring 2021 29th Mediterranean Conference on Control and Automation, MED 2021, 2021, pp. 173–178, 9480180.

[7] Casalino, G., Castellano, G., Kaymak, U., **Zaza, G.** Balancing Accuracy and Interpretability through Neuro-Fuzzy Models for Cardiovascular Risk Assessment. *2021 IEEE Symposium Series on Computational Intelligence, SSCI 2021 - Proceedings*, 2021.

[8] Casalino, G., Castellano, G., **Zaza, G.** A mHealth solution for contact-less self-monitoring of blood oxygen saturation. *Proceedings - IEEE Symposium on Computers and Communications*, 2020, 2020-July, 9219718.

[9] Casalino, G., Castellano, G., Pasquadibisceglie, V., **Zaza, G.** Contact-less real-time monitoring of cardiovascular risk using video imaging and fuzzy inference rules. *Information (Switzerland)*, 2019, 10(1), 9.

[10] Castellano, G., Castiello, C., Pasquadibisceglie, V., **Zaza, G.** FISDeT: Fuzzy inference system development tool. *International Journal of Computational Intelligence Systems*, 2017, 10(1), pp. 13–22.