

Sustainability of Symbiotic AI

Teacher

Graziella De Martino

Course description (min 150, max 300 words)

As Symbiotic AI systems become more popular, it's important to consider the high costs associated with data analysis and gathering. This includes both computational and human resources. To ensure sustainability, it's essential to address both the environmental and economic impact of these solutions. Luckily, new AI techniques can help with this by allowing for the reuse of data and models without sacrificing accuracy. This course is focused on reducing computational resources and minimizing the ecological impact of AI-based modeling. This will be done from two perspectives:

1. Reducing the effort required for data gathering, labeling, and representation; and
2. Recycling models to save computational resources for model learning and tuning.

Course period

December 2023 – January 2024

SSD

INF/01

Course References

- [1] Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. <https://doi.org/10.48550/ARXIV.1906.02243>
- [2] Anthony, L. F. W., Kanding, B., & Selvan, R. (2020). Carbontracker : Tracking and Predicting the Carbon Footprint of Training Deep Learning Models (arXiv:2007.03051). arXiv. <http://arxiv.org/abs/2007.03051>
- [3] Tan, M., & Le, Q. V. (2020). EfficientNet : Rethinking Model Scaling for Convolutional Neural Networks (arXiv:1905.11946). arXiv. <http://arxiv.org/abs/1905.11946>
- [4] Jaureguiualzo, E. (2011). PUE: The Green Grid metric for evaluating the energy efficiency in DC (Data Center). Measurement method using the power demand. 2011 IEEE 33rd International Telecommunications Energy Conference (INTELEC), 1-8. <https://doi.org/10.1109/INTELEC.2011.6099718>
- [5] Hernandez, D., & Brown, T. B. (2020). Measuring the Algorithmic Efficiency of Neural Networks (arXiv:2005.04305). arXiv. <http://arxiv.org/abs/2005.04305>

Credits and Hours

3 credits, two of lectures (16 hours) and one of practice (15 hour), for a total of 31 hours.

Exam Modality

To pass the exam, students have two options available to them.

1. The first option is to give a **paper presentation** where they present the contents of two papers recommended by the teacher. It is important to note that group presentations are not allowed.

2. The second option is to work on a **project** where the student implements and experimentally validates an algorithm or its variation from a paper recommended by the teacher. Projects can be done individually or in groups of up to three students, depending on the algorithm.

Teacher(s) CV

Please find my CV attached at the end of this document.

Teacher(s) Main Publications

Graziella De Martino, Gianvito Pio, Michelangelo Ceci (2023). Multi-View Overlapping Clustering for the Identification of the Subject Matter of Legal Judgments, Information Sciences, DOI: 10.1016/j.ins.2023.118956.

Graziella De Martino, Gianvito Pio (2022). Identification of Paragraph Regularities in Legal Judgements Through Clustering and Textual Embedding. In: Ceci, M., Flesca, S., Masciari, E., Manco, G., Raś, Z.W. (eds) Foundations of Intelligent Systems. ISMIS 2022. Lecture Notes in Computer Science, vol 13515. Springer, Cham, DOI: 10.1007/978-3-031-16564-1_8.

Graziella De Martino, Gianvito Pio, Michelangelo Ceci (2021). A Two-Step Method based on Embedding and Clustering to Identify Regularities in Legal Case Judgements (Discussion Paper). Proceedings of the 29th Italian Symposium on Advanced Database Systems, {SEBD} 2021, Pizzo Calabro (VV), Italy, <http://ceur-ws.org/Vol-2994/paper42.pdf>, September 5-9, 2021.

Graziella De Martino, Gianvito Pio, Michelangelo Ceci, PRILJ: An Efficient Two-Step Method Based On Embedding And Clustering For The Identification Of Regularities In Legal Case Judgments, Artificial Intelligence And Law, DOI: 10.1007/S10506-021-09297-1, 2021.

Lorenzo de Mattei, **Graziella De Martino**, Andrea Iovine, Alessio Miaschi, Marco Polignano, Giulia Rambelli, ATE_ABSITA @ EVALITA2020: Overview of the Aspect Term Extraction and Aspect-based Sentiment Analysis Task. EVALITA 2020, DOI: 10.4000/books.aaccademia.6849

GRAZIELLA DE MARTINO

PERSONAL INFORMATION

Email: de.martino.1991@gmail.com

Contact number: (+39) 3318088576

Residential address: Via Re David, 67, Bari, BA, 70125, Italy

EDUCATION

**University of Bari “Aldo Moro”,
Italy**

2019 – 2022

**Doctor of Philosophy in Computer Science and Mathematics [2019 Scholarship
Winner]**

Research Project Title: The Use of Machine Learning and Process Discovery to Support
Legal Transcript Writing

Key Research Areas: Legal Informatics, Digital Transformation, Machine Learning, Natural
Language Processing, Word and Document Embedding, Process Discovery.

Overall Classification of the Qualification: Pass with Honors and Doctor Europaeus
certification.

Utrecht University, Holland

March 2022 – June 2022

Research Internship

Developing a multi-view classification system for legal case judgments capable of capturing
both the semantics of the textual content and the co-citations of legal acts and precedents.

**European University Institute,
Italy & University of Pittsburg
School of Law, US**

8 – 16 July 2021

Artificial Intelligence and Law Summer School [online]

Key Research Areas: Introduction to Artificial Intelligence and Law, Knowledge
Representation, Rule-based Systems, Modelling Legal Arguments, Case-base Reasoning,
Evidential Reasoning, and Machine Learning and Legal Analytics.

Middlesex University, UK

2016 – 2019

Master of Science in Business Information Systems Management

Thesis Title: Leveraging Digital Disruption Through the Implementation of a Decision
Support System for Civil Courts

Overall Classification of the Qualification: Second class (Upper Division)

**Malta College of Arts, Science
and Technology, Malta**

2013 – 2015

Bachelor of Science in Computer Networks

Thesis Title: Routing Load Balancing Using an Optimised Ant Colony Algorithm

Overall Classification of the Qualification: Second class (Upper Division)

WORK EXPERIENCE

**University of Bari “Aldo Moro”,
Italy**

January 2023 – present

Assistant Professor (non-tenure track) in Computer Science

Research Project: FAIR – Future Artificial Intelligence Research, an Italian project approved
by the National Recovery and Resilience Plan (NRRP), that defines the theoretical, modelling
and engineering aspects of modern Artificial Intelligence.

Main Area of Focus: Designing, implementing, and validating new Machine Learning
techniques for model recycling.

**University of Bari “Aldo Moro”,
Italy**

November 2022 – January 2023

Research Assistant

Research Project: Next Generation UPP, which the Ministry of Justice funds with the
contribution of the European Union, National Operational Program Governance, and
Institutional Capacity 2014-2020.

Main Area of Focus: Developing new collaborative schemes between universities and judicial
offices to improve the efficiency of justice administration in Southern Italy.

DESCRIPTION OF RESEARCH ACTIVITIES

My area of research revolves around the intersection of law and technology. I am primarily focused on enhancing **Legal Information Retrieval, Embedding, Clustering,** and **Approximate Nearest Neighbor Search** techniques. To achieve this goal, I have developed methods that enable a more nuanced understanding of legal rulings and the grouping of similar cases.

I have also explored the application of **Natural Language Processing** to enhance the interpretability of legal documents. Additionally, I have delved into **explainable AI** approaches to ensure transparent decision-making in sensitive legal contexts. In one of my workshop publications, I applied **Sentiment Analysis** techniques to evaluate opinions and product reviews.

PUBLICATIONS

- Journal Paper: [P1]**
26 September 2022
- Graziella De Martino,** Gianvito Pio, Michelangelo Ceci (2023). Multi-View Overlapping Clustering for the Identification of the Subject Matter of Legal Judgments, Information Sciences, DOI: 10.1016/j.ins.2023.118956.
- Conference Proceedings: [P2]**
26 September 2022
- Graziella De Martino,** Gianvito Pio (2022). Identification of Paragraph Regularities in Legal Judgements Through Clustering and Textual Embedding. In: Ceci, M., Flesca, S., Masciari, E., Manco, G., Raś, Z.W. (eds) Foundations of Intelligent Systems. ISMIS 2022. Lecture Notes in Computer Science, vol 13515. Springer, Cham, DOI: 10.1007/978-3-031-16564-1_8.
- Discussion Paper: [P3]**
1 November 2021
- Graziella De Martino,** Gianvito Pio, Michelangelo Ceci (2021). A Two-Step Method based on Embedding and Clustering to Identify Regularities in Legal Case Judgements (Discussion Paper). Proceedings of the 29th Italian Symposium on Advanced Database Systems, {SEBD} 2021, Pizzo Calabro (VV), Italy, <http://ceur-ws.org/Vol-2994/paper42.pdf>, September 5-9, 2021.
- Journal Paper: [P4]**
4 August 2021
- Graziella De Martino,** Gianvito Pio, Michelangelo Ceci, PRILJ: An Efficient Two-Step Method Based On Embedding And Clustering For The Identification Of Regularities In Legal Case Judgments, Artificial Intelligence And Law, DOI: 10.1007/S10506-021-09297-1, 2021.
- Workshop Proceedings: [P5]**
17 December 2020
- Lorenzo de Mattei, **Graziella De Martino,** Andrea Iovine, Alessio Miaschi, Marco Polignano, Giulia Rambelli, ATE_ABSITA @ EVALITA2020: Overview of the Aspect Term Extraction and Aspect-based Sentiment Analysis Task. EVALITA 2020, DOI: 10.4000/books.aaccademia.6849.

PRESENTATIONS AS A SPEAKER AT CONFERENCES OR WORKSHOPS

- ITAL-IA 2022: CINI National Conference on Artificial Intelligence, Italy**
9 – 11 February 2022
- Presentation Title:** Machine learning methods to support tasks in the legal field.
- SEBD 2021: Italian Symposium on Advanced Database Systems, Italy**
5 – 9 September 2021
- Presentation Title:** A Two-Step Method based on Embedding and Clustering to Identify Regularities in Legal Case Judgements (Discussion Paper).

PARTICIPATION AS ORGANISATION COMMITTEE MEMBER

- NFMCP 2023 @ ECML-PKDD: The 11th edition of the Workshop New Frontiers in Mining Complex Patterns, Italy**
18 – 22 September 2023
- Role:** Organizing chair.
<http://www.di.uniba.it/mignone/NFMCP23/index.html>

ITADATA 2023: 2nd Italian Conference on Big Data and Data Science, Italy
11 – 13 September 2023

Role: Hackathon chair.
<https://www.itadata.it/2023/committees>

DeepLearn 2023: 9th International School on Deep Learning, Italy
3 – 7 April 2023

Role: Local co-chair.
<https://deeplearn.irdta.eu/2023sp/>

PROJECT ACTIVITIES

University of Bari “Aldo Moro”, Italy
December 2022

Research Project: IMPETUS - Intelligent Management of Processes, Ethics and Technology for Urban Safety
Main Area of Focus: Using neural network-based techniques to identify anomalies in data.

University of Bari “Aldo Moro”, Italy
October 2022

Research Project: CounteR – Privacy-First Situational Awareness Platform for Violent Terrorism and Crime Prediction, Counter Radicalization and Citizen Protection
Main Area of Focus: Using embedding techniques to classify textual messages in social media.

TEACHING ACTIVITIES

University of Bari “Aldo Moro”, Italy
January 2020 – 2022

Computational and Complexity Theory – Bachelor of Science in Computer Science.

University of Bari “Aldo Moro”, Italy
January 2020 – 2021

Computer Science Laboratory – Bachelor of Science in Computer Science and Digital Communication.

PERSONAL SKILLS

Technical Skills

Database management systems – **Oracle Database**, and **MySQL**. Solid technical background with a complete understanding of database design and modelling.

Programming languages – **Python** and **R**. Solid technical background with a complete understanding of existing machine learning algorithms and libraries (examples include but are not limited to nltk, scikit-learn, gensim, torch, stanza, networkx and pm4py). Expertise in implementing new machine learning methods.

Graphic editors – **Adobe Photoshop CS6**, and **Adobe InDesign**.

Document preparation systems – **MS Office** and **LaTeX**.

Languages

Maltese (Native)
English (High proficiency)
Italian (High proficiency)
French (Elementary proficiency)