VI International Conference on Optimization, Learning Algorithms and Applications

Book of Abstracts

Ada Byron - Málaga, Spain

June, 10 - 12, 2026

VI International Conference on Optimization, Learning Algorithms and Applications

OL2A 2026

Book of Abstracts

ol2a.ipb.pt

Editors

Ana I. Pereira, Polytechnic Institute of Bragança, Portugal

Ismael Navas, University of Málaga, Spain

Florbela P. Fernandes, Polytechnic Institute of Bragança, Portugal

João Paulo Coelho, Polytechnic Institute of Bragança, Portugal

João P. Teixeira, Polytechnic Institute of Bragança, Portugal

José Lima, Polytechnic Institute of Bragança, Portugal

Maria F. Pacheco, Polytechnic Institute of Bragança, Portugal

Rui Pedro Lopes, Polytechnic Institute of Bragança, Portugal

Instituto Politécnico de Bragança

Campus de Santa Apolónia

5300-253 Bragança - Portugal

ISBN: 978-972-745-373-3

http://hdl.handle.net/10198/36787

Contents

Welcome

It is our pleasure to welcome you to OL2A 2026 – the International Conference on Optimization, Learning Algorithms, and Applications.

OL2A has established itself as an international forum for researchers, academics, and professionals working in the fields of optimization and learning, promoting the exchange of ideas, the presentation of recent scientific advances, and the development of new collaborations. Now in its sixth consecutive edition, the conference continues to bring together contributions covering a broad range of topics, including multi-objective optimization, machine learning approaches for optimization, decision-making under uncertainty, and industrial applications.

In addition to the technical sessions and scientific discussions, all the contributions presented at the conference will be published in the conference proceedings as a volume in the Springer book series, further enhancing the dissemination and impact of the research presented at OL2A 2026.

The Organizing Committee is pleased to welcome all participants to OL2A 2026 and hopes that this event will provide a stimulating environment for scientific exchange, networking, and the establishment of future research partnerships.

The OL2A 2026 organization committee,
Ana I. Pereira, Ismael Navas, Florbela P. Fernandes, João P. Coelho, João P. Teixeira, José Lima, Maria F. Pacheco, and Rui Pedro Lopes

Committees

General Chairs

Ana I. Pereira, Polytechnic Institute of Bragança, Portugal
Ismael Navas, University of Málaga, Spain

Local Team Committee

Cristóbal Barba, University of Málaga, Spain
Sandro Hurtado, University of Málaga, Spain
Maciej Ribinski, University of Málaga, Spain

Program Committee Chairs

Florbela P. Fernandes, Polytechnic Institute of Bragança, Portugal
João Paulo Teixeira, Polytechnic Institute of Bragança, Portugal
Maria F. Pacheco, Polytechnic Institute of Bragança, Portugal
Rui Pedro Lopes, Polytechnic Institute of Bragança, Portugal

Technology Chairs

João Paulo Coelho, Polytechnic Institute of Bragança, Portugal
José Lima, Polytechnic Institute of Bragança, Portugal

Scientific Committee

Abdulghafom Mohammad, University of West, Sweden

Abeer Alsadoon, Charles Sturt University, Australia

Absalom Ezugwu, North-West University, South Africa

Ademir Aparecido Constantino, State University of Maringá, Brazil

Ala' Khalifeh, German Jordanian University, Jordan

Alberto Nakano, Federal University of Technology of Paraná, Brazil

Alexandre Douplik, Ryerson University, Canada

Alfonso González-Briones, University of Salamanca, Spain

Aline Miotto Amaral, State University of Maringá, Brazil

Ana Flávia P. de Camargos, CEFET - Rio de Janeiro, Brazil

Ana I. Pereira, Polytechnic Institute of Bragança, Portugal

Ana Maria A. C. Rocha, University of Minho, Portugal

Ana Paula Teixeira, University of Trás-os-Montes and Alto Douro, Portugal

Anderson Faustino da Silva, State University of Maringá, Brazil

André Pinz Borges, Federal University of Technology of Paraná, Brazil

Andrej Košir, University of Ljubljana, Slovenia

Ângela Silva, University of Minho, Portugal

António José Sánchez-Salmerón, Polytechnic University of Valencia, Spain

António Valente, University of Trás-os-Montes and Alto Douro, Portugal

Armando Mendes, Azores University, Portugal

Arnaldo Cândido Júnior, Federal University of Technology of Paraná, Brazil

B. Rajesh Kanna, Vellore Institute of Technology, India

Beatriz F. Azevedo, Polytechnic Institute of Bragança, Portugal

Benjamín J. González Díaz, University of La Laguna, Spain

Bilal Ahmad, University of Warwick, United Kingdom

C. Sweetlin Hemalatha, Vellore Institute of Technology, India

Carlos Henrique Alves, CEFET - Rio de Janeiro, Brazil

Carmen Galé, University of Zaragoza, Spain

Carolina Gil Marcelino, Federal University of Rio de Janeiro, Brazil

Christopher Expósito-Izquierdo, University of La Laguna, Spain

Clara Bento Vaz, Polytechnic Institute of Bragança, Portugal

Damir Vrancic, Jozef Stefan Institute, Slovenia

Davide Anguita, University of Genoa, Italy

Diamantino Silva Freitas, University of Porto, Portugal

Diego Brandão, CEFET - Rio de Janeiro, Brazil

Eduardo Vinicius Kuhn, Federal University of Technology of Paraná, Brazil

Elaine Mosconi, Université de Sherbrooke, Canada

Eligius M. T. Hendrix, University of Malaga, Spain

Eris Zeqo, Aleksandër Moisiu University of Durrës, Albania

Fatiha Djebar, Abdulghafour Mohammad, University of West

Fatmira Prodani, Aleksandër Moisiu University of Durrës, Albania

Felipe Nascimento Martins, Hanze University of Applied Sciences, Netherlands

Fernando Dela Prieta Pintado, University of Salamanca, Spain

Filipe Alves, University of Minho, Portugal

Florbela P. Fernandes, Polytechnic Institute of Bragança, Portugal

Francisco Sedano, University of León, Spain

Fredrik Danielsson, University of West, Sweden

Frida Gjermeni, Aleksandër Moisiu University of Durrës, Albania

Gediminas Daukšys, Kauno Technikos Kolegija, Lithuania

Gislaine Lapasini Leal, State University of Maringá, Brazil

Giulio Ferro, University of Genoa, Italy

Glaucia Maria Bressan, Federal University of Technology of Paraná, Brazil

Heba M. Afify, Cairo University, Egypt

Helena Sofia Rodrigues, Polytechnic Institute of Viana do Castelo, Portugal

Ismael Delgado, University of Malaga, Spain

J. Marcos Moreno, University of La Laguna, Spain

Jasminka Hasic, Sarajevo School of Science and Tecnology, Bosnia and Herzegovina

Joanna Kolodziej, Cracow University of Technology, Poland

João Paulo Coelho, Polytechnic Institute of Bragança, Portugal

João P. Teixeira, Polytechnic Institute of Bragança, Portugal

João Paulo Carmo, University of São Paulo, Brazil

Jorge Garcia, Polytechnic Institute of Viana do Castelo, Portugal

Jorge Igual, Polytechnic University of Valencia, Spain

Jorge Ribeiro, Polytechnic Institute of Viana do Castelo, Portugal

José Boaventura-Cunha, University of Trás-os-Montes and Alto Douro, Portugal

José Cascalho, Azores University, Portugal

José Lima, Polytechnic Institute of Bragança, Portugal

Joseane Pontes, Federal University of Technology of Paraná, Brazil

Josip Music, Faculty of Electrical Eng., Mechanical Eng. and Naval Architecture

Juan A. Méndez, University of La Laguna, Spain

Juan Alberto García Esteban, University of Salamanca, Spain

Juani L. Redondo, University of Almeria, Spain

Júlio Cesar Nievola, Pontifícia Universidade Católica do Paraná, Brazil

Karine Deschinkel, Université Marie & Louis Pasteur, France

Kristina Sutiene, Kaunas University of Technology, Lithuania

Lidia Sánchez, University of León, Spain

Lindita Mukli, Aleksandër Moisiu University of Durrës, Albania

Lino Costa, University of Minho, Portugal

Luca Oneto, University of Genoa, Italy

Luca Spalazzi, Marche Polytechnical University, Italy

Luis Antonio De Santa-Eulalia, Université de Sherbrooke, Canada

Luís Coelho, University of Porto, Portugal

Maria F. Pacheco, Polytechnic Institute of Bragança, Portugal

Mahmood Reza Khabbazi, University of West, Sweden

Manuel Castejón Limas, University of León, Spain

Marc Jungers, University of Lorraine, France

Marco Aurélio Wehrmeister, Federal University of Technology of Paraná, Brazil

Marek Nowakowski, Military Institute of Armoured and Automotive Technology in Sulejowek, Poland

Maria João Pereira, Polytechnic Institute of Bragança, Portugal

Maria Rigkou, University of Patras, Greece

Martin Hering-Bertram, Hochschule Bremen, Germany

Mattias Bennulf, University of West, Sweden

Mauro Henrique Mulati, State University of Maringá, Brazil

Michael Paraskevas, University of Peloponnese, Greece

Michal Podpora, Opole University of Technology, Poland

Michela Robba, University of Genoa, Italy

Miguel Ángel Prada, University of León, Spain

Milena Pinto, CEFET - Rio de Janeiro, Brazil

Miroslav Kulich, Czech Technical University Prague, Czech Republic

Ouamaya Baala Canalda, Université Marie & Louis Pasteur, France

Paulo Alves, Polytechnic Institute of Bragança, Portugal

Paulo Leitão, Polytechnic Institute of Bragança, Portugal

Paulo Lopes dos Santos, University of Porto, Portugal

Paulo Moura Oliveira, University of Trás-os-Montes and Alto Douro, Portugal
Pavel Pakshin, Nizhny Novgorod State Tech University, Russia
Pedro João Rodrigues, Polythecnic Institute of Bragança, Portugal
Pedro Luiz de Paula Filho, Federal University of Technology of Paraná, Brazil
Pedro Miguel Rodrigues, Catholic University of Portugal, Portugal
Pedro Pinto, Polytechnic Institute of Porto, Portugal
Philippe Canalda, Université Marie & Louis Pasteur, France
Raphaël Couturier, Université Marie & Louis Pasteur, France
Roberto Molina de Souza, Federal University of Technology of Paraná, Brazil
Roberto Tavares, Federal University of São Carlos, Brazil
Rodrigo Thom de Souza, State University of Maringá, Brazil
Rui Pedro Lopes, Polytechnic Institute of Bragança, Portugal
Sabrina Łuman, University of Applied Sciences of Rijeka, Croatia
Sancho Salcedo Sanz, University of Alcalá, Spain
Sandro Ridella, University of Genoa, Italy
Sani R. Silva, Federal University of Technology of Paraná, Brazil
Santiago Torres, University of La Laguna, Spain
Saptadeep Biswas, GITAM Visakhapatnam, India
Sara Paiva, Polytechnic Institute of Viana do Castelo, Portugal
Shridhar Devamane, Global Academy of Technology - Karnataka, India
Slawomir Stepien, Poznan University of Technology, Poland
Spiros Sirmakessis, University of Peloponnese, Greece
Sudha Ramasamy, University of West, Sweden
Teresa Perdicoulis, University of Trás-os-Montes and Alto Douro, Portugal
Uta Bohnebeck, Hochschule Bremen, Germany
Virginia Castillo, University of León, Spain
Vítor Duarte dos Santos, Nova University Lisbon, Portugal
Vítor Hugo Pinto, University of Porto, Portugal
Wojciech Giernacki, Poznan University of Technology, Poland
Wojciech Paszke, University of Reims Champagne-Ardenne, France
Wynand Alkema, Hanze University of Applied Sciences, Netherlands
Zahia Guessoum, University of Reims Champagne-Ardenne, France

Invited Plenary Lectures

From Dialogue to Decision: Orchestrating Agents for Conversational Data Exploration and Multi-Objective Optimization

Sihem Amer-Yahia

Biography: Sihem Amer-Yahia is a Silver Medal CNRS Research Director and Deputy Director of the Lab of Informatics of Grenoble. She works on exploratory data analysis and algorithmic upskilling. Prior to that she was Principal Scientist at QCRI, Senior Scientist at Yahoo! Research and Member of Technical Staff at at&t Labs. Sihem served as PC chair for SIGMOD 2023 and as the coordinator of the Diversity, Equity and Inclusion initiative for the database community. In 2024, she received the 2024 IEEE TCDE Impact Award, the SIGMOD Contributions Award, and the VLDB Women in Database Award.

Abstract: Modern scientific discovery and decision-making necessitate agents capable of navigating both complex datasets and multifaceted, often conflicting, goal landscapes. This talk explores the intersection of conversational data exploration, an incremental process where agents help users articulate needs through data interaction, and multi-objective sequential decision making. While traditional Reinforcement Learning (RL) has advanced task-specific agent training, we are witnessing a paradigm shift toward zero-shot orchestration. By leveraging a spectrum of techniques, from specialized RL policies to general-purpose Large Language Models (LLMs), we can reuse existing single-objective policies to solve complex multi-objective problems without training from scratch. We will discuss empirical evidence from the field of Education showing that this orchestrated approach can achieve competitive Pareto quality while reducing computational cost. The talk concludes by examining open research questions regarding how LLM context richness and reflective foresight enable agents to bridge the gap between incremental data-driven insights and optimal sequential actions.

To trust or not to trust, is that the only question? A journey from safety to responsibility in AI systems

Javier Del Ser

Biography: Javier (Javi) Del Ser holds a Ph.D. in Control Engineering and Industrial Electronics from the University of Navarra (2006, Cum Laude), and a second Ph.D. in Information and Communication Technologies from the University of Alcalá de Henares (2013, Cum Laude and recipient of the Extraordinary PhD Award). He is a Principal Investigator in Applied Artificial Intelligence and Chief AI Scientist at TECNALIA. Additionally, he is a Distinguished Researcher at the Department of Mathematics of the University of the Basque Country (UPV/EHU), and a Visiting Professor at the University of Granada (Spain) and the University of Natural Resources and Life Sciences, Vienna (Austria). His research focuses on Artificial Intelligence, Machine Learning and Evolutionary Computation, with applications in practical modeling and optimization challenges across diverse sectors, including industry, healthcare, transportation, energy, and mobility. He has coauthored more than 480 scientific papers, edited 7 books, supervised 20 doctoral theses, and contributed to more than 70 research projects and industrial contracts. He has been included in the Highly Cited Researchers (HCR) list by Clarivate, reflecting the global influence and impact of his scientific work, with a yearly rate of more than 10,000 citations to his authored works (2025). He is a Senior Member of IEEE and a recipient of several awards for his research trajectory, including the BRTA Award (among more than 3,000 researchers in the Basque R&D network).

Abstract: Artificial Intelligence is rapidly evolving from foundation models to pervasive, autonomous decision-making systems. In this technological upsurge, the question of trustworthiness in AI has become central to public, industrial, and scientific discourse. But trustworthiness, while essential, is only one piece of a much broader challenge. This keynote explores the complex landscape that spans AI safety, trustworthy AI, and the emerging need for responsibility. We begin by examining the foundations of AI safety: robustness, reliability, and alignment with human values. Building on this, we delve into the principles of trustworthy AI, including transparency, fairness, and privacy, and why these properties are increasingly difficult to guarantee in modern AI systems characterized by scale, data-driven behaviors, and opaque internal dynamics. The talk will then confront the realities of today's AI momentum, where unprecedented model availability, unrestricted usage, and high social impact converge with insufficient governance,

unclear accountability, and escalating risks. From misinformation to autonomous decision failures, the potential consequences of poorly supervised AI deployment are severe and far-reaching. Ultimately, we argue that trust and safety, while necessary, are no longer sufficient. The field must move toward responsible AI: a comprehensive ecosystem of frameworks, standards, and shared practices that ensure AI systems are not only capable and trustworthy, but also governed, monitored, and deployed with societal well-being at their core. The keynote will conclude with a call to action for researchers, practitioners, and institutions to embrace responsible AI as the guiding paradigm for the next generation of intelligent AI-based systems.

Special Sessions

AIHaM - Artificial Intelligence in Healthcare and Medicine

Organizers:

Beatriz Flamia Azevedo

Polytechnic Institute of Bragança, Portugal

beatrizflamia@ipb.pt

Spiros Sirmakessis

University of Peloponnese, Greece

syrma@uop.gr

Description: Artificial intelligence (AI), optimization, and data analytics significantly impact healthcare systems. Data-driven diagnostic systems and machine learning algorithms are important tools in healthcare to help professionals identify, analyze, and understand patient conditions and disease development. These systems enable early detection and intervention, which are crucial to reducing disease-related complications and supporting decision-making. Therefore, research on AI innovations and challenges in healthcare and medicine is welcome.

Topics: Optimization · Machine Learning · Case Studies · Data Analysis · Healthcare · AI for diagnostics · Patient data privacy and security.

AI4SAgE - Artificial Intelligence for Sustainable Agriculture and Environment

Organizers:

Juan Albino Mendez Perez

University of La Laguna, Spain

jamendez@ull.es

Eduardo Bezerra da Silva

CEFET - Rio de Janeiro, Brazil

ebezerra@cefet-rj.br

Rui Pedro Lopes

Polytechnic Institute of Bragança, Portugal

rlopes@ipb.pt

Description: The session Artificial Intelligence for Sustainable Agriculture and Environment (AI4SAgE) aims to explore the transformative role of AI in addressing critical environmental and agricultural challenges. As the world faces increasing pressure from climate change, biodiversity loss, and the need for sustainable food production, AI-based approaches offer new possibilities for predictive modeling, resource optimization, environmental monitoring, and decision support. This session welcomes interdisciplinary contributions that bridge AI methods including machine learning, computer vision, remote sensing, and decision systems with real-world applications in agriculture, forestry, water management, and ecological preservation. We invite both theoretical advances and practical implementations, particularly those that promote sustainability, resilience, and ethical considerations in the use of intelligent systems for environmental and agricultural innovation.

Topics: AI for Sustainable Agriculture · Environmental Monitoring and Climate Resilience · Remote Sensing and Earth Observation · Ethics, Equity, and Responsible AI in Agri-Environmental Systems.

SODDD - Statistics and Optimization for Data-Driven Decisions

Organizers:

Helena Sofia Rodrigues

ADiT-Lab, Instituto Politécnico de Viana do Castelo, and Center for Research & Development in Mathematics and Applications, University of Aveiro, Portugal

sofiarodrigues@esce.ipvc.pt

Ângela Silva

Algoritmi R&D Center, University of Minho, Portugal

asilva@dps.uminho.pt

Jorge Esparteiro Garcia

ADiT-Lab, Instituto Politécnico de Viana do Castelo, Portugal

jorgegarcia@esce.ipvc.pt

Description: Contemporary decision-making problems increasingly arise in complex environments characterized by large volumes of data, uncertainty, and multiple, often conflicting objectives. In such contexts, the integration of Statistical Learning and Optimization provides a rigorous and complementary framework for extracting knowledge from data and translating it into effective and actionable decisions. Statistical learning methods enable the modeling, prediction, and quantification of uncertainty from heterogeneous data sources, while optimization techniques support the identification of efficient and robust solutions under operational, economic, and societal constraints. This special session aims to bring together researchers and practitioners working at the intersection of applied statistics, machine learning, and optimization, with applications spanning areas such as healthcare, energy systems, transportation, industry, environmental management, and public policy. The session welcomes methodological contributions as well as real-world applications demonstrating how data-driven modeling and optimization can support informed decision-making in complex applied settings, fostering interdisciplinary collaboration and advancing quantitative approaches to contemporary decision challenges.

Topics: Optimization · Data.

Abstracts

Is Prediction Enough? A Dynamic Graph Attention and Reinforcement Learning Framework for Automated Trading

Dhruv Jain, and Sobhan Sarkar

IIM Ranchi, India

IIT Kharagpur, India

Stock market prediction is inherently complicated because of the nature of the financial data, which are both stochastic and nonlinear. While recent graph-based approaches successfully exploit corporate relations, they suffer from two key limitations: (i) reliance on static knowledge graphs that ignore temporal market dynamics and (ii) a focus on prediction that cannot be directly translated into actionable portfolio management strategies to make money. To address these issues, we propose DART (Dynamic Attention & Reinforcement Trading), a framework that integrates dynamic graph neural networks with deep reinforcement learning. Using historical price data extracted from Yahoo Finance, DART constructs time-varying adjacency matrices from sliding-window price correlations, enabling the model to capture short-term, evolving inter-stock dependencies. The node feature space is further enriched with technical indicators, resulting in a more expressive representation of the market state. Instead of stopping at movement classification, DART uses a Proximal Policy Optimization (PPO) agent to allocate long-only portfolio weights, directly optimizing a risk-aware reward that accounts for returns, volatility, drawdown, and transaction costs. Experiments demonstrate DART outperforms state-of-the-art baselines, achieving a 2.24 Sharpe ratio and 245% cumulative return with a significantly lower maximum drawdown of 3.316%. These results highlight DART's effectiveness in bridging the gap between predictive modeling and actionable, risk-aware trading strategies.

Keywords: Dynamic Relational Graphs · Graph Neural Networks · Deep Reinforcement Learning · Stock Market Prediction

A Constrained DBSCAN Algorithm for Insect Facility Location under Data-Scarce Conditions

Safa Vakili, Maria João Santos, Jaime S. Cardoso, and Mariana Carvalho

CIICESI, ESTG, Polytechnic of Porto, Portugal

INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal

Insect farming for alternative protein production requires strategic site selection, yet data scarcity often hinders optimal facility placement. This study proposes a data-driven approach for locating an insect production facility when only limited supplier information is available. The approach relies on density-based clustering (DBSCAN), combined with nutrient composition constraints, and estimates the facility location using a weighted center of mass of eligible suppliers. It was applied to a dataset consisting of 110 suppliers and their feedstock across Portugal, with materials categorized into protein, carbohydrate, and complementary groups.The resulting clustering identified a single group representing 49.7% of the available feedstock, whose combined nutrient profile satisfies the predefined criteria for insect feed formulation. The weighted centroid of this cluster suggests an optimal facility location, particularly under a spatial distance threshold of 35 km for supplier inclusion. The sensitivity analysis indicates that the choice of spatial threshold influences both the geographic concentration of clusters and the range of materials included. Tighter thresholds produce more compact clusters, whereas higher thresholds expand material coverage.Furthermore, 78.16% of suppliers excluded as noise or outliers were primarily supplement providers, underscoring the importance of nutrient composition as a key criterion for supplier selection.These findings demonstrate that integrating nutrient-based constraints with density clustering can effectively guide site selection for insect production in data-scarce scenarios, balancing logistical feasibility with nutritional adequacy.

Keywords: DBSCAN · Facility location Problem · Insect production · Circular economy · Composition constraints.

A Multinomial Logistic Regression Approach to Investigating Breast Cancer Molecular Subtypes

Ana Beatriz Valentin, Glaucia Bressan, and Elisângela Lizzi

Federal University of Technology of Paraná (UTFPR), Brazil

Breast cancer is a highly heterogeneous disease whose molecular subtypes are associated with distinct biological behaviors, therapeutic responses, and clinical outcomes. In this study, factors associated with breast cancer molecular subtypes are investigated, using a multinomial logistic regression framework applied to clinical and genomic data from The Cancer Genome Atlas (TCGA). The outcome variable comprised five molecular subtypes (Luminal A, Luminal B, Basal, HER2, and Normal-like), while predictors included categorical and numerical variables. After that, a Directed Acyclic Graph (DAG) was constructed to support causal interpretability, guiding the specification of univariable and multivariable models. Crude and adjusted Odds Ratios with 95% confidence intervals were estimated, using Luminal A as the reference category. The results highlight TMB as a key genomic factor, showing strong and subtype-specific associations after multivariable adjustment, including marked suppression effects observed for the Basal subtype. Age at diagnosis acted as an important confounder, influencing both subtype distribution and clinical outcomes. Clinical variables such as pathological staging and survival status exhibited significant associations only after appropriate adjustment, underscoring the importance of multivariable modeling. Model adequacy was supported by goodness-of-fit diagnostics, residual analyses, likelihood ratio tests, and cross-validation, indicating satisfactory discrimination and calibration. Overall, this study demonstrates that carefully structured multinomial models, informed by DAG-based reasoning, provide an interpretable and statistically robust framework for investigating breast cancer molecular subtypes and their associated clinical and genomic profiles.

Keywords: Breast cancer molecular subtypes · Directed acyclic graphs · Clinical-genomic integration · Multinomial logistic regression

Promoting Simple Agents: Ensemble Methods for Event-Log Prediction

Benedikt Bollig, Matthias Függer, Thomas Nowak, and Paul Zeinaty

Université Paris-Saclay, CNRS, ENS Paris-Saclay, LMF, Gif-sur-Yvette, France

We compare lightweight automata-based models (n-grams) with neural architectures (LSTM, Transformer) for next-activity prediction in streaming event logs. Experiments on synthetic patterns and five real-world process mining datasets show that n-grams with appropriate context windows achieve comparable accuracy to neural models while requiring substantially fewer resources. Unlike windowed neural architectures, which show unstable performance patterns, n-grams provide stable and consistent accuracy. While we demonstrate that classical ensemble methods like voting improve n-gram performance, they require running many agents in parallel during inference, increasing memory consumption and latency. We propose an ensemble method, the promotion algorithm, that dynamically selects between two active models during inference, reducing overhead compared to classical voting schemes. On real-world datasets, these ensembles match or exceed the accuracy of non-windowed neural models with lower computational cost.

Keywords: Event-log prediction · N-gram · LSTM · Transformer · Ensemble methods · Agents

A Machine Learning Workflow for Biomedical Tabular Data Based on Exploratory Feature Analysis

Jamile de Souza, Glaucia Bressan, Fabrício Martins Lopes, and Ulisses da Rocha

Federal University of Technology of Paraná (UTFPR), Brazil

Helmholtz-Centre for Environmental Research, Germany

Machine learning has become a key tool in biomedical research, enabling advances in diagnosis, prognosis, and clinical decision making. However, biomedical tabular datasets often present challenges such as high dimensionality, heterogeneous feature types and class imbalance. Despite these complexities, many applied studies prioritise predictive performance, while exploratory data analysis and feature selection procedures are under-reported or implicitly described, limiting model interpretability. This study proposes a structured machine learning workflow for labelled tabular data that explicitly emphasises exploratory analysis and feature selection as central methodological steps rather than auxiliary preprocessing tasks. The proposed approach is evaluated using a publicly available Alzheimer's disease dataset as a representative case study. As a result, exploratory analyses identified five features with consistent discriminative potential across visual, statistical, and probabilistic criteria, while also revealing partial inter-class overlap and low feature redundancy. These findings directly informed subsequent modelling decisions and guided the application of multiple supervised classification algorithms. As a result, the supervised classification models achieved balanced accuracies of up to 95%, while preserving interpretability and reducing computational complexity through feature reduction. The features identified during the exploratory stage were consistent with those ranked as the most influential by the best-performing predictive model, reinforcing the coherence between exploratory insights and predictive outcomes.

Keywords: Exploratory Analysis · Feature Engineering · Model Interpretability · Methodological Framework · Pattern Recognition

Multi-Model Document Layout Analysis Based on YOLO with OCR Fusion for Structured Newspaper Digitization

Paulo Roberto Machado Silva Junior, Paulo Alves, José E. Fernandes, Bruno F. Silva, Davide E. Dias, and Marcio Cunha

Federal University of Technology of Paraná (UTFPR), Brazil

Polytechnic Institute of Bragança, Portugal

This paper examines the limitations of conventional Optical Character Recognition (OCR) when applied to documents with diverse layouts, such as newspapers and magazines, where the process can transcribe individual characters but often fails to retain structural organization and logical reading order, leading to disordered text. To address these issues, the proposed project integrates Document Layout Analysis (DLA) with OCR in a multi-stage process, where YOLOv10 and YOLOv12 models detect and segment document elements, and the resulting regions are then passed to PaddleOCR for text extraction. Experimental results show that the first pre-trained model achieved a mAP@50 of 0.728 on a 2,000 images sample from DocLayNet, while the second pretrained model achieved a mAP@50 of 0.519 on a custom dataset. The fusion approach reduced detection redundancy, and comparative evaluation against a production process indicates good performance. The final process creates a JSON output that keeps the association between the bounding box coordinates and the extracted text. Future work will suggest the use of vision language models (VLMs) to improve reading order and improve the post-processing approach.

Keywords: Document Layout Analysis · Optical Character Recognition · PaddleOCR

A Continuous-Time Formulation for the Unrelated Parallel Machine Scheduling Problem with Setup Times and Additional Resources in the Setups

Mario Villaizán-Vallelado, Matteo Salvatori, Carlos Segura, and Belén Carro Martinez

Telefónica Scientific Research, Spain

Universidad de Valladolid, Spain

Production planning in environments with unrelated parallel machines, sequence-dependent setup times and scarce auxiliary resources can be formalized as the Unrelated Parallel Machine scheduling problem with setup times and additional resources in the setups (UPMSR-S), an NP-hard problem that arises in several industrial contexts where setup operations consume limited resources such as skilled operators or specialized equipment. We introduce a new mathematical formulation for the UPMSR-S that significantly reduces the number of decision variables while preserving the essential interactions between machine assignment, job sequencing, setup times, and resource availability. By compactly modelling resource-dependent setup feasibility, the proposed formulation improves scalability and provides a more efficient basis for solution methods. Computational experiments on a large benchmark set show that the new formulation leads to substantial reductions in model size and computational effort when compared with existing formulations, while maintaining solution quality. These results highlight the effectiveness of the proposed formulation for tackling large and difficult scheduling instances, and demonstrate its potential as a robust modelling framework for advanced solution approaches in production environments with setup and resource constraints.

Keywords: Scheduling · Unrelated Parallel Machine · Additional resources · Continuous-Time · Formulation

Sustainable Clean-in-Place Process through Digital Twin-Guided Optimization

Guido Parodi, Elisa Anzolin, Francesca Cipollini, Cecilia Muraro, Paolo Favella, Samuele Pozzacchio, Nicola Pretto, Daniele Bonventre, Gianmarco Sabbatini, Alberto Chiarello, and Luca Oneto

ECOR, Italy

University of Genoa, Italy

aizoOn consulting, Italy

Cleaning Validation is a fundamental and rigorously documented requirement in the pharmaceutical industry; It provides reproducible evidence that product residues and cleaning agents are reduced below scientifically established levels. The complex Clean-in-Place process, which combines mechanical action, chemical agents, heat, and time, is challenging to optimize due to the large number of influential variables and the dynamic nature of the underlying phenomena. To address this challenge, a *Digital Twin* was developed to model the intricate cause-effect relationships leading to the desired cleaning outcomes. The model employs Decision Trees because of their effectiveness in the small sample setting and interpretability in critical applications compared with other state-of-the-art models. The resulting Digital Twin provides the constraints for a mathematical optimization problem designed to minimize sustainability key performance indicators, such as the equivalent CO_2 emissions associated with water consumption, heating energy, and chemical usage, while ensuring compliance with the defined cleanliness requirements. Results obtained from a real-world case study confirm the validity and practical applicability of the proposed approach.

Keywords: Cleaning in Place · Setting Optimization · Digital Twin

Detecting and Mitigating Overvalidation in Adaptive Data Analysis: an Information-Theoretic Approach

Fabrizio Mori, Antonio Emanuele Cinà, Davide Anguita, Fabio Roli, and Luca Oneto

University of Genoa, Italy

Over the past decades, machine learning has achieved remarkable improvements in performance.When applied to real-world problems, these advances are often accompanied by the need to optimize numerous design choices, including the selection of algorithms, optimization methods, model architectures, and hyperparameters. To reduce the computational burden as well as the selection and estimation bias of this process, adaptive techniques have been developed to efficiently, selectively, and effectively explore the search space. In this work, we investigate the emergence of overvalidation, namely selection and/or estimation bias arising from excessive adaptivity to validation data, and we propose strategies to mitigate its effects. To this end, we first conduct a series of experiments across multiple tasks, datasets, and algorithms to empirically demonstrate the existence of this phenomenon. We then leverage a state-of-the-art information-theoretic framework from statistical learning theory, i.e., the mutual-information-based generalization bound for adaptive data analysis, to provide theoretical insights into the phenomenon and to inform principled approaches for its detection and mitigation.

Keywords: Overvalidation · Adaptive Data Analysis · Selection Bias · Estimation Bias · Information Theory

Automated Norberg Angle Measurement: A Regression Network Approach with Weighted Angular Loss

Tânia Gonçalves, Lio Gonçalves, Pedro Franco, Mário Ginja, and Vítor Filipe

University of Trás-os-Montes and Alto Douro (UTAD), Portugal

Canine Hip Dysplasia (CHD) diagnosis involves a multiparametric radiographic evaluation. The Norberg Angle (NA) is a quantitative metric for assessing joint laxity, however, its manual measurement is prone to inter-observer variability. Standard deep learning (DL) approaches often treat landmark detection as a purely spatial task, ignoring underlying anatomical relationships. This study proposes an anatomical consistency hybrid keypoint-regression framework to automate NA measurement and CHD classification. We evolved a regression baseline into a hybrid system by integrating a Weighted Angular Loss function, which penalizes geometric inconsistencies during training. The method was validated on a test set of 114 x-rays (228 rotationally aligned hip joints) using a ResNet50V2 backbone and benchmarked against DenseNet121, VGG16 and MobileNetV2. The proposed approach achieved a Mean Angular Error of 2.93° and a Specificity of 98.2% in binary classification. The model demonstrated a high discriminative capability with an AUC of 0.95, confirming its consistency even in borderline cases. These results suggest that constraining deep learning models with biomechanical logic significantly enhances diagnostic reliability, offering a safe screening tool for veterinary breeding programs.

Keywords: Canine Hip Dysplasia · Deep Learning · Keypoint Regression · Geometric Constraints · Norberg Angle

Surrogate-assisted optimization with adaptive hypervolume-controlled fidelity scheduling for renewable energy community design

Philipp Thunshirn, José Baptista, and Tiago Pinto

University of Applied Sciences Technikum Wien, Austria

University of Trás-os-Montes and Alto Douro (UTAD), CPES - INESCTEC, Portugal

This paper proposes an adaptive multi-fidelity optimization (AMFO) framework that combines a high-fidelity optimization model with the NSGA-II and a random forest-based surrogate model. The AMFO controller dynamically allocates evaluation budget across model fidelities based on the hypervolume error. Several mechanisms are employed to prevent surrogate drift and increase the number of higher-fidelity evaluations when the hypervolume exceeds a predefined threshold level. The framework is demonstrated for the bi-objective optimization of a renewable energy community with a highly expensive building and stochastic vehicle-to-community (V2C) model based on hourly time-series data. The surrogate predicts lifetime-aggregated energy flows, from which objectives are calculated deterministically. Results show that AMFO reduces runtime from 404 minutes (high-fidelity baseline) to 8.4 minutes (-98%) with only 1.6% of high-fidelity evaluations

compared to surrogate-only optimization (7.7 minutes, 0 truth evaluations), AMFO improves Pareto-front fidelity, increasing holdout metrics from $R^2 = 0.978$ to 0.998 while reducing rel. MAE from 2.30% to 0.5% and from 1.77% to 0.4% respectively. The AMFO controller is particularly effective for nonlinear objectives, as adaptive high-fidelity corrections prevent surrogate drift and maintain global Pareto-front accuracy.

Keywords: Multi-fidelity optimization · Surrogate modeling · Energy communities · Random Forest

Detection and Quantification of the Morgan Line in Veterinary Radiographs: A Hybrid Approach Combining Intensity Histogram Extraction and Deep Learning

Filipa Alves, Vítor Filipe, Pedro Franco-Gonçalo, Mário Ginja, Antônio Lacreta, Sofia Alves-Pimenta, Bruno Colaço, and Lio Gonçalves

Animal and Veterinary Research Centre (CECAV), University of Trás-os-Montes and Alto Douro (UTAD), Portugal

Department of Engineering, University of Trás-os-Montes and Alto Douro (UTAD), Portugal

Department of Veterinary Science, University of Trás-os-Montes and Alto Douro (UTAD), Portugal

Universidade Federal de Lavras, Lavras, Brazil

Canine hip dysplasia (CHD) is a prevalent orthopedic condition in large-breed dogs, where early radiographic detection of the caudolateral curvilinear osteophyte (CCO), or Morgan line, remains challenging due to its subtle morphology and inter-observer variability. While deep learning (DL) could standardize this process, its application is often limited by scarce annotated veterinary datasets. To address this bottleneck, this study proposes a hybrid DL pipeline. The approach extracts 1D intensity histograms (IH) from the femoral neck region in 263 standard ventrodorsal hip-extended (VDHE) radiographs. A Multi-task neural network (MultiTaskNet) was trained to simultaneously detect the presence of the osteophyte and quantify its position and thickness. On stratified test sets, the model achieved exceptional binary classification (BC) performance, with an area under the curve (AUC) of 99.88%, accuracy of 98.62%, and F1-score of 97.99%. For the regression tasks, it recorded a mean absolute error (MAE) of 2.02 pixels for position and 1.39 pixels for thickness. Clinical validation demonstrated substantial agreement with expert consensus (Kappa = 0.81, 95% CI [0.73, 0.88]). However, Bland-Altman analysis revealed a systematic underestimation in advanced cases (bias 0.61 mm). Despite this limitation in continuous measurement, the model excels at the primary BC task. These findings validate IH-augmented multi-task learning as a highly effective, lightweight alternative to standard convolutional neural network (CNN) for CHD screening in data-constrained settings.

Keywords: Canine Hip Dysplasia · Caudolateral curvilinear osteophyte · Deep Learning · Intensity histograms

Integration of Deep Learning and Diffusion Models for the Detection of Gastric Intestinal Metaplasia in Narrow-Band Imaging

Fabio Andrade, Sílvio Capela, Jorge Lage, and Vítor Filipe

Department of Engineering, University of Trás-os-Montes and Alto Douro (UTAD), Portugal

Serviço de Gastrenterologia, Unidade Local de Saúde de Trás-os-Montes e Alto Douro (ULSTMAD), Portugal

Gastric Intestinal Metaplasia (GIM) is a critical precancerous condition in the progression of gastric cancer, with Narrow-Band Imaging (NBI) technology being superior to white light for its diagnosis. Despite, the development of robust Deep Learning systems for GIM detection they are severely hindered by the critical scarcity and high imbalance of annotated medical datasets, which stands as a major bottleneck. To overcome this data deficit, this study proposes a generative pipeline utilizing Stable Diffusion fine-tuned via LoRA to synthesize high-fidelity NBI images. The methodology integrates artifact removal through LaMa inpainting and a rigorous contour-based quality control to ensure anatomical plausibility, validated by specific metrics such as MedFID and MS-SSIM. The performance of five classifiers (including Custom Convolutional Neural Networks (CNN), pre-trained architectures, and Transformers) were evaluated under real, classical augmentation, and hybrid data strategies. Results demonstrate that the hybrid strategy consistently outperformed the approaches, with DenseNet121 achieving an accuracy of 94.2%. Notably, the Custom CNN trained from scratch registered the highest relative improvement, proving that synthetic data effectively fills the information gaps in limited datasets. We conclude that high-fidelity generative augmentation is a viable solution to the data scarcity challenge, significantly enhancing diagnostic robustness.

Keywords: AI · Data Scarcity · Synthetic Data Augmentation

Facial Expression Recognition with CNNs and Vision Transformers: Full Images, Landmarks, and a Hybrid Method on FER-2013, FER+ and RAF-DB

Daniel Casanova, Alessandra B. G. Hoffmann, Arnaldo Candido Junior, Pedro João Rodrigues, and Pedro Luiz de Paula Filho

Research Centre in Digitalization and Intelligent Robotics (CeDRI), SusTEC, Instituto Politécnico de Bragança, Portugal

Department of Computer Science, Federal University of Technology of Paraná (UTFPR), Brazil

Instituto de Biociências, Letras e Ciências Exatas, São Paulo State University (UNESP), São José do Rio Preto, Brazil

Facial Expression Recognition (FER) is an important task in computer vision, but its deployment in real-world scenarios remains challenging due to data variability, illumination changes, pose variations, occlusions, and privacy concerns. This study presents a controlled comparison of three FER input modalities: full-face images, facial landmarks extracted using MediaPipe and dlib, and a mixed fusion strategy combining both representations. Two backbone architectures, ResNet-50 and Vision Transformers (ViT), are evaluated under a unified experimental pipeline. Experiments are conducted on three benchmark datasets: FER-2013, FER+, and RAF-DB. Performance is assessed using Macro F1-score, accuracy, and per-image inference time, along with analyses of landmark density and extractor choice. Results indicate that full-image ViT achieves the highest recognition performance, while ResNet-50 offers lower inference latency. Landmark-based inputs improve efficiency and privacy at the cost of moderate performance loss, and the fusion strategy provides a balanced trade-off.

Keywords: Facial Expression Recognition · Facial Landmarks · Privacy-preserving Methods · Vision Transformer · ResNet · Deep Learning

Generative Diffusion Models for Probabilistic Electric-Vehicle Charging Session Forecasting

Manasi Jeste, and Zeliha Kamaci

Fraunhofer-Institut für Solare Energiesysteme ISE, Germany

Smart charging systems for electric vehicles (EVs) require accurate and reliable forecasts of charging behavior to reduce waiting times, improve charger utilization, and coordinate operations under power-grid constraints. In practical charging infrastructures, vehicle arrivals are observed at plug-in time, while key quantities for planning departure time, connection duration, and delivered energy remain uncertain. We propose a station-level probabilistic forecasting framework based on Generative AI using conditional diffusion models with a Transformer denoiser. Using only historical session timestamps and lightweight calendar context, the method generates coherent probabilistic predictions of departure times, durations, and per-session energy without relying on user identifiers, battery specifications, or state-of-charge measurements. Unlike independent quantile regression, the diffusion formulation models the joint conditional distribution of session characteristics, preserving consistency between arrival, departure, duration, and energy. Experiments on a public Commercial DC fast-charging dataset show competitive point accuracy and improved uncertainty calibration, particularly under congestion. We further propagate the probabilistic forecasts to operational KPIs such as station occupancy and energy demand, supporting risk-aware planning under uncertainty.

Keywords: EV charging · Generative AI · Probabilistic forecasting · Diffusion models · Charging session forecasting · Smart charging

Motivations, Perceived Impacts, and Brand Discovery on TikTok: A Quantitative Study of Portuguese Young Adults

Cristiana Silva, Inês Fernandes, Alison Gonçalves, and Helena Sofia Rodrigues

ADiT-Lab, Instituto Politécnico de Viana do Castelo, Portugal

Instituto Politécnico de Viana do Castelo, Portugal

TikTok has become a prominent digital platform among young adults, influencing patterns of media consumption, social interaction, and brand discovery. Despite its growing relevance, quantitative evidence on users' motivations, perceived impacts, and consumer-related behaviours remains limited in smaller European contexts. This study presents a quantitative analysis of TikTok use among Portuguese young adults, focusing on motivations for use, perceived impacts, brand and product discovery, and future perceptions of the platform. Data were collected through an online survey administered to 134 respondents aged 18 - 29 years, of whom 117 active TikTok users were retained for analysis. Descriptive statistics, reliability analysis, principal component analysis, synthetic index construction, and cluster analysis were employed. Results indicate that entertainment is the dominant motivation for TikTok use, while brand and product discovery constitutes a statistically relevant but secondary dimension. Although users acknowledge TikTok's role in enhancing product awareness, its direct influence on purchasing decisions remains moderate. Cluster analysis reveals two distinct user profiles: a highly engaged group with stronger psychological and commercial susceptibility, and a lower-engagement group with limited perceived impact. Overall, the findings provide structured quantitative evidence on TikTok usage behaviours among young adults and offer empirical insights relevant to digital consumer behaviour research and applied modelling.

Keywords: TikTok · Young Adults · Social Media Use · Motivations · Perceived impacts · Brand Discovery · Principal Component Analysis · Cluster analysis

Enhancing a Niche Algorithm with Local Search: a Multimodal Optimization Approach

Hugo Nascimento Silva, Beatriz Flamia Azevedo, and Carolina Gil Marcelino

Federal University of Rio de Janeiro, Brazil

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

This study aims to improve the performance of the Fitness Euclidean-distance Ratio Particle Swarm Optimization (FER-PSO), a niche algorithm, by modifying and implementing local search operations and sub-swarm guidance using k-means clustering to solve multimodal optimization problems. The proposed approach aims to reduce the need for a large population to find the global optima of multimodal functions. To this end, some modifications were incorporated into the calculation of the scalar factor and the equation of movement, and local search strategies were implemented using the modified Hill Climbing algorithm. Experiments were conducted on multimodal test functions from the literature. The results showed that the modifications can reduce the number of particles required and increase the algorithm's efficiency at identifying global minima, especially for highly complex functions.

Keywords: PSO · Hill Climbing · Multimodal Optmization · Clustering

Forecasting Television Audiences: A Comparative Study of Machine Learning Models

Bruno F. Silva, Davide E. Dias, Paulo Alves, and José E. Fernandes

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Accurate forecasting of television audiences is essential for broadcast scheduling, advertising optimization, and resource allocation in increasingly competitive media environments. While most prior studies rely heavily on autoregressive signals derived from past audience ratings, this work investigates the feasibility of predicting television viewership exclusively using exogenous variables known prior to broadcast. This study conducts a comparative evaluation of linear regression, Ridge regression, Random Forest, LightGBM, and Long Short-Term Memory (LSTM) networks using large-scale audience data from four Portuguese television channels spanning 2021-2025. The feature set includes temporal, seasonal, and program-typology variables engineered to capture cyclical viewing patterns without incorporating historical audience values. Experimental results across multiple temporal regimes demonstrate that ensemble-based models consistently outperform linear approaches, achieving coefficients of determination above 0.93. Deep learning models show competitive performance but exhibit greater sensitivity to data volume and temporal segmentation. The findings indicate that high predictive accuracy can be achieved without endogenous audience inputs, highlighting the structural importance of broadcast timing and content characteristics. These results contribute to the design of operational decision support systems for television programming and advertising planning.

Keywords: Television Audience Forecasting · Time Series Prediction · Random Forest · LightGBM · LSTM Network · Ensemble Learning

Algorithmic vs. Manual Packing in Warehouse Picking: A Case Study

Ângelo Soares, António Galrão Ramos, Pedro Rocha, Miguel Sousa, and Carlos Maio

INESC TEC, ISEP, Polytechnic of Porto, Portugal

Luís Simões, Portugal

This paper investigates a real-world pallet loading problem arising from the automation of picking operations at a logistics service provider. The challenge involves determining how items should be consolidated into stacks and how these stacks are allocated and positioned on pallets, while accounting for their subsequent effect on picker routing. Beyond conventional packing restrictions, the problem imposes an additional constraint requiring that all stacks associated with the same product type be assigned to a single pallet. The main objective is to minimize the total number of pallets used, while a secondary goal considers the reduction of warehouse travel distances during the picking process. We introduce a novel algorithm, PackStack, designed specifically to optimize pallet utilization. While the algorithm does not directly tackle the routing problem, the study assesses the impact of its packing solutions on warehouse travel distances. To estimate picker travel distances, heuristic methods are employed that approximate realistic warehouse movements with low computational effort. Computational experiments on 52 real-world instances indicate that Pack-Stack never produces a worse solution than the current process, and in 69% of the instances, it produces solutions with a smaller number of pallets, reaching an overall reduction of about 22% in pallet usage. When examining the effect on travel distances, the proposed packing solutions lead to shorter distances in just over half of the instances (56%), while around 35% of cases experience longer routes. In addition, the packing-focused strategy increases total travel distance by about 7%, reflecting the inherent trade-off between pallet efficiency and routing performance.

Keywords: PackStack · Packing · Routing · Picking

An Empirical Study of Ensemble Diversity in AutoML for Regression

Salomey Osei, Andrés R. Masegosa, and Antonio D. Masegosa

Aalborg University, Denmark

University of Deusto, Spain

Ensemble-based Automated Machine Learning (AutoML) methods have become increasingly prominent for regression tasks due to their ability to combine diverse predictive models to improve generalization performance. Despite their empirical success, the mechanisms underlying this performance, particularly the role of model diversity, remain insufficiently explored. In this study, we leverage recent theoretical frameworks on ensemble diversity to investigate how diversity influences regression performance in AutoML systems. We focus on AUTO-SKLEARN, a widely used ensemble-based AutoML framework, and examine how individual model diversity and predictive performance evolve across its four key phases: Base Learners, Meta-Learning, Bayesian Optimization (BO), and Caruana Ensemble. Using 13 diverse datasets, we empirically quantify these effects using Mean Squared Error (MSE) and diversity measures. Our analysis reveals a clear trade-off between individual model accuracy and ensemble diversity. Although the BO phase selects high-performing individual regressors, the aggregated mean MSE across datasets remains comparable across phases (0.0342 in Base and 0.0372 in BO). This can be attributed to dataset heterogeneity rather than lack of optimization effect. Additionally, it does not consistently increase ensemble diversity relative to earlier phases (from 0.2475 in Base to 0.2656 in BO). In contrast, the Caruana phase selects models with slightly higher MSE (0.0297) but significantly increases ensemble diversity (0.2882), resulting in better overall generalization. These findings show that diversity is a key driver of ensemble performance in regression tasks in AutoML methods. Overall, this work provides both theoretical and empirical evidence that diversity is critical to the success of regression in AutoML. By quantifying how each phase contributes to the trade-off between MSE and diversity, we provide insights into ensemble construction strategies to inform the development of more robust AutoML frameworks for regression tasks.

Keywords: Automated Machine Learning (AutoML) · AUTO-SKLEARN · Bayesian Optimization (BO) · Diversity · Ensemble Learning · Regression

AutoXML, AI-based Automatic Annotation of Documents

Tomás Araújo, Maria João Varanda Pereira, and Pedro Henriques

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Universidade do Minho, Portugal

Digital annotation of documents is important because it turns unstructured text into structured, machine-interpretable data, enabling automation, analysis, and reuse. Doing it manually, however, is time consuming, error-prone, and hard to scale. This article is focused on the design and development of an automatic document XML annotation tool based on the Document Type Definition (DTD). AutoXML, available at github, is a Python package with a simple interface, working as a tool to manage annotation projects. It allows creating annotation projects and editing documents in XML, based on a DTD, including two different approaches for automatic annotation, one based on more traditional NLP pipelines, and the other leveraging the more recent LLM. At the end, the results of both approaches are compared, allowing us to draw some conclusions about the novelty of using LLM for this purpose.

Keywords: Document Annotation · XML Markup · DTD · NLP · LLM

Impact of Synthetic Data Composition on Deep Learning Models for Robotic Visual Inspection

Johann Amorim, Milena Faria Pinto, Gabriel Araujo, Accácio Ferreira dos Santos Neto, Yuri Nascimento Silva, and José Lima

CEFET/MG, Brazil

CEFET/RJ, Brazil

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

The performance of deep learning models in robotic visual inspection strongly depends on the realism and distribution of the training data. In solar farm inspection scenarios, acquiring annotated real-world datasets is costly and operationally restrictive. To mitigate this limitation, this work investigates two controlled synthetic data sources: (i) real-synthetic images, defined as photorealistic composite images created by overlaying target objects onto artificial backgrounds, and (ii) simulated-synthetic images, generated in the Gazebo simulation environment integrated with ROS. Using the YOLOv5m architecture, five training configurations were evaluated: 100% Gazebo simulated data, 100% real-synthetic data, and three mixed compositions (25/75, 50/50, and 75/25). The objective was to quantify how dataset composition influences cross-domain generalization between simulated and photorealistic domains. Results demonstrate that training exclusively on Gazebo data yields superior performance in simulated environments (mAP50 = 0.944, mAP50-95 = 0.829), but limited generalization to photorealistic real-synthetic images. Conversely, increasing the proportion of real-synthetic images significantly improves real-domain performance. The best trade-off was achieved with 75% real-synthetic and 25% Gazebo data, reaching mAP50 = 0.480 and mAP50-95 = 0.269 on real-synthetic test data. These findings confirm that photorealistic real-synthetic images are critical for improving domain transfer, while Gazebo-generated data remains essential for controlled simulation validation.

Keywords: Robotic visual inspection · Photorealistic real-synthetic images · Deep learning · Solar farm inspection

Biparametric Biobjective Linear Programming

Kezang Yuden, and Stefan Ruzika

RPTU Kaiseslautern-Landau, Germany

Multi-parametric linear programming and multi-objective linear programming are both well researched fields of optimization, and we explore the connection between these two fields. In this article, we consider parametric linear programs with multiple objective functions that depend linearly on multiple parameters. Using the weighted sum scalarization of a multi-objective linear program, we find efficient solutions of the biparametric biobjective linear program. We develop a theoretical framework to solve parametric biobjective problems with two parameters and generalize the framework to a multi-parametric biobjective linear program. This framework relates the structure of the weight set decomposition of the associated multi-objective linear program to the parameter set of the biparametric linear program.

Keywords: Multi-objective optimization · Parametric programming · Weighted sum scalarization

Exploiting delivery zone structure for last-mile route sequencing: a decision tree approach

António Galrão Ramos, Manuel Lopes, and Pedro Rocha

INESC TEC, ISEP, Polytechnic of Porto, Portugal

Last-mile routing is one of the biggest challenges supply chain managers face today. Despite important advances in solving routing problems, a significant gap remains between theoretical route planning and real-world route execution, mainly because route quality depends on a multitude of factors related to geography, infrastructure, and customers that are hard to address using classical optimization methods. This article presents the decision tree-based algorithm designed by Team MEGI for the Amazon Last Mile Routing Research Challenge, which aimed to incorporate the tacit knowledge of drivers into routing algorithms. The resulting application was ranked 4th out of 45 finalist teams. The key components are a decision tree and an innovative sequencing engine that exploits the Amazon districting strategy. We also present an improved version that significantly outperforms the competition submission.

Keywords: Last-mile routing problem · Driver learning · Tacit knowledge

RAG for Veterinary Pharmaceuticals: Evaluating Local LLMs in Portuguese

Guilherme Quintero Lorenzi, Maria João Varanda Pereira, and Gleifer Vaz Alves

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Federal University of Technology of Paraná - Campus Ponta Grossa, Brasil

A common assumption in large language model (LLM) deployment is that larger models consistently outperform smaller ones in both accuracy and efficiency. We challenge this assumption through systematic evaluation of 8 local open-source LLMs (8B–14B parameters) across 4 model families on Portuguese veterinary-pharmaceutical information retrieval tasks. We present an automated pipeline integrating web scraping, PDF parsing, and Retrieval-Augmented Generation (RAG) to extract and organize domain-specific knowledge from unstructured regulatory sources, enabling specialized query answering without reliance on commercial APIs. Evaluating on 35 real-world queries from a pharmaceutical distributor, we find that model architecture and training quality dominate parameter count: a well-selected 9B model outperforms a 20B model by 34.2 percentage points, and scaling within families is non-linear, with some newer/larger models underperforming predecessors. These findings have direct implications for resource-constrained deployment scenarios, showing that careful model selection from mid-size options (8B–9B) can achieve superior accuracy-latency trade-offs compared to indiscriminate use of larger models.

Keywords: Retrieval-Augmented Generation · Large Language Models · Model Scaling · Domain-Specific Information Retrieval · Veterinary Pharmaceuticals

Development of a Digital Twin Platform for Agriculture 4.0 Demonstration

João Paulo Coelho, Romina Eliana Giugge, and J. A. B. Coelho

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

This paper presents the development of a LoRa-based digital twin platform designed as an Agriculture 4.0 demonstrator. The proposed system integrates distributed IoT sensing nodes, LoRaWAN communication, edge computing, and FIWARE-based context management within a unified cyber-physical architecture. Interoperability and scalability are enabled by creating a semantic data model compliant with NGSI-LD standards. The digital twin core relies on a process-based crop growth model based on classical balance equations. This model continuously consumes real-time environmental measurements, such as temperature, radiation, and CO_2 concentration, to predict biomass evolution. Experimental validation demonstrates reliable data acquisition, real-time context synchronisation, and consistent qualitative behaviour of the growth model under varying environmental conditions. The proposed demonstrator provides an accessible, scalable platform for illustrating digital twin principles in agriculture, bridging the gap between advanced IoT infrastructure and practical greenhouse management applications.

Keywords: Digital Twin · Agriculture 4.0 · LoRaWAN · FIWARE · Smart Greenhouse · Internet of Things

Mitigating Catastrophic Forgetting in Lifelong Knowledge Editing via Persistent Adam

Timur Turushev, and Maria Khodorchenko

ITMO University, Russia

Large Language Models (LLMs) store a substantial amount of factual knowledge in their parameters, but updating this knowledge via retraining is costly. Knowledge editing (KE) aims to efficiently rewrite specific facts while preserving unrelated behavior. In practice, existing KE methods struggle to balance edit success, locality, and generalization, and sequential editing further introduces catastrophic forgetting, where later edits degrade earlier ones. In this work, we build on the WISE lifelong model editing framework and focus on an overlooked component of gradient-based editing: the optimizer state. We propose persistent-state Adam, a variant that carries moment estimates across consecutive fact edits instead of resetting them, thereby accumulating a per-parameter importance signal that dampens updates to frequently modified weights. We evaluate on ZsRE and WikiBig using three instruction-tuned LLMs (Qwen2.5-7B, Llama-3.1-8B, Mistral-7B) with up to 1,000 sequential edits. Across most studied WISE settings, persistent Adam improves reliability and generalization, with the strongest gains appearing on ZsRE and at longer edit horizons, and yields up to a 17% gain in the mean of reliability, generalization, and locality in the best studied settings. Ablations show that the benefit stems almost entirely from persisting the second moment, while persisting only the first moment can be harmful; a comparison with learning-rate warmup confirms that the gains reflect genuinely useful cross-edit information rather than mere early-step stabilization. Our findings suggest that preserving optimizer state can materially improve lifelong knowledge editing without changing the core editing mechanism or model architecture.

Keywords: Knowledge editing · Lifelong editing · Catastrophic forgetting · WISE · Adam

Optimizing Student Roommate Compatibility Using a Genetic Algorithm Hybridized with Reinforcement Learning

Matteo Johnson, Gabe Hardy-Fowler, Kevin Keopp, Eric Matage, and Zulal Sevkli

Miami University, United States

This study proposes a solution to optimizing dorm room assignments by refining a combination of genetic algorithm properties with a reinforcement learning-based agent. Given a set of students with individual living habits and characteristics, the goal of this research is to develop an algorithm that maximizes each student's satisfaction by assigning them the most suitable roommate. Preliminary research includes reports on room-assignment-based problems using genetic algorithms and hybrid forms of genetic algorithms. Data were collected using a technique inspired by this research, gathering synthetic student data, which consisting of an ID string followed by characteristics denoting the student's hobbies, registered major, career cluster, and preferences. Weighted scores are created for each student characteristic, and a satisfaction score is calculated based on student room assignments. Later, a reinforcement learning agent is implemented in the mutation phase of the algorithm to enhance new candidate assignments. Based on the results of the final algorithm, reinforcement learning combined with a genetic algorithm approach displays a significant improvement in overall student satisfaction over the baseline genetic algorithm's results, achieving a 43.4% improvement in peak fitness.

Keywords: Genetic Algorithm · Reinforcement Learning · Dormitory Assignment Problem · Reinforce Algorithm · Machine Learning

Comparative Evaluation of Large Language Models for Synthetic Data Generation: A Multidimensional Analysis of Quality, Diversity, and Privacy

Mohamed Zohir Koufi, Youcef Moulai, and Zahia Guessoum

CAPGEMINI, France

Universty of Reims Champagne Ardenne, France

The generation of synthetic data has emerged as a critical solution for addressing privacy regulations and data scarcity in application testing and development. This paper presents a comparative evaluation of three state-of-the-art Large Language Models (LLMs), Meta Llama 3 8B, Mistral 7B, and EXAONE 3.5 32B, for synthetic data generation. We introduce a multidimensional evaluation framework encompassing statistical fidelity (KL divergence, Wasserstein distance), structural preservation (correlation matrices, mutual information), diversity metrics (unique-value ratio, Jaccard similarity), and privacy protection measures (Jensen-Shannon divergence, overfitting detection). Experiments conducted on two datasets of varying complexity (ANIME and EMPLOYEE) reveal significant performance differences across models and data types. EXAONE 3.5 achieved the best legitimate overall balance with a normalized weighted global score of 0.626 on a [0, 1] scale, while Mistral 7B demonstrated critical overfitting on tabular data by reproducing training data exactly. Meta Llama 3 excelled in semantic understanding but exhibited higher statistical divergence and categorical instability. Our findings emphasize the necessity of context-specific model selection and rigorous privacy auditing before deployment, with computational efficiency varying by $3.3x$ across models.

Keywords: GenAI · LLM · Synthetic Data Generation

Optimal Crew Selection in Competition using Multiobjective Optimization Techniques.

Jonathan Pereira Mendoza, and Christian Fernández-Campusano

Universidad de Alicante, Spain

Universidad de Santiago de Chile USACH, Chile

In rowing, crew composition is critical for maximizing performance and ensuring vessel stability. This study addresses the problem of selecting an optimal crew for a Valencian llaut, a traditional boat that requires a balanced and efficient configuration. The objective is to develop an optimization model to determine the ideal eight rowers from a pool of candidates, considering factors such as power, weight, and each rower's position preferences, while maintaining the boat's lateral balance and other sport-specific constraints.

Keywords: Linear programming · Multi-objective optimization · Team selection · Integer programming · Sports analytics · Rowing performance · Crew balance

CompLiTT: An End-to-End Pipeline for Processing and Summarizing Historical Books

Luis Resende, Rui Pedro Lopes, and Alexia Bravo

Instituto Politécnico de Bragança, Portugal

Automatic summarization of historical texts faces significant structural and linguistic challenges that precede the application of language models, including the physical degradation of source materials, the use of archaic typographies, and non-standardized orthography. This work presents CompLiTT, an integrated and replicable pipeline designed for the automated processing and summarization of historical books dating from the 16th century onward. The proposed methodology employs the PaddleOCR framework for text extraction in Markdown format, demonstrating effectiveness in preserving semantic layout and mitigating digitization noise. The study applies a hierarchical summarization strategy to a multilingual dataset composed of 26 works in five languages (Portuguese, English, French, Italian, and Spanish). The effectiveness of the pipeline was evaluated through cross-validation across different Large Language Models (LLMs) using the BooookScore, ROUGE, and BERTScore metrics. The results show that the *gpt-5-mini* model achieved superior performance in generating clear and comprehensive summaries, although high semantic consistency was observed among all evaluated models. This work contributes to the field of digital humanities by establishing a methodological foundation for knowledge retrieval in historical collections.

Keywords: Automatic Summarization · Historical Documents · Large Language Models

Experimental Evaluation of Thermal Cameras for Fiducial Marker-Based UAV Operations

Guido S. Berger, João Braun, Arezki Chellal, Luciano Bonzatto, Matheus Ferraz, João Mendes, Ana I. Pereira, António Valente, and José Lima

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

University of Trás-os-Montes and Alto Douro (UTAD), Portugal

Federal University of Technology of Paraná (UTFPR), Brazil

Universität Osnabrück, Germany

This work presents a methodology for characterizing two thermal camera models applied for the identification of a thermal fiducial marker. The study analyzed the marker's detection performance in three distinct scenarios: a controlled environment, an outdoor environment, and an airborne condition, using a multirotor Unmanned Aerial Vehicle (UAV). The methodology considered the analysis of translational and rotational errors, angular sensitivity, and the behavior of depth estimates based on the Z-component of the thermal marker's pose. The results indicated that both sensors maintained detection capability even under adverse conditions, including direct sunlight exposure. In a controlled environment, both models showed low dispersion among the evaluated metrics, demonstrating experimental consistency and predictable behavior of the pose estimates. Experiments conducted in outdoor and airborne environments indicate greater sample robustness for the Seek Thermal PRO thermal camera model. In contrast, the XR model demonstrated greater stability and less distance dependence in depth estimates. The results indicate the future applicability of thermal cameras for navigation and autonomous landing operations by multirotor UAVs in outdoor environments subject to lighting variations, which usually compromise systems based on visible-spectrum cameras, highlighting their potential for integration into architectures based on sensor fusion techniques

Keywords: Thermal Camera · Sensor Characterization · Fiducial Marker Detection · Unmanned Aerial Vehicle

Multi-UAV Path Planning for Forest Surveillance based on ACO-BmTSP algorithm

Sílvia Pereira, Eduardo Solteiro Pires, and Paulo Moura Oliveira

Instituto Politécnico de Bragança, Portugal

University of Trás-os-Montes and Alto Douro (UTAD), Portugal

Forest surveillance is fundamental for environmental monitoring and for reducing the risk of fires, requiring efficient coverage of large and complex areas. Systems that utilize multiple Unmanned Aerial Vehicles (UAVs) have emerged as a promising alternative, due to their adaptability and scalability. However, coordinating multiple UAVs while considering coverage efficiency and battery constraints remains a challenging optimization problem. This paper proposes a multi-UAV forest inspection approach based on an Ant Colony Optimization algorithm adapted from the Balanced Multiple Traveling Salesman Problem (ACO-BmTSP). The monitoring model is performed using a cover path planning problem, where discrete inspection points represent sensing locations in a continuous forest area. The proposed approach aims to optimize UAV route allocation and integrates UAV battery limitations and coverage quality into the optimization process. The performance of the proposed method is tested in a real-world scenario using a well-known Portuguese forest in the center of the country, the Leiria pine forest. The performance of the method is evaluated through computational experiments and shows that it is capable of generating efficient and balanced multi-UAVs inspection routes, achieving full area coverage while maintaining feasible energy consumption levels for all UAVs. The proposed framework demonstrates good scalability and provides a flexible optimization tool for large-scale autonomous forest monitoring applications.

Keywords: Multiple traveling salesman problem · Cover path planning · UAV · ACO algorithm

A Cyber-Physical Water Sampling System for Distributed Wastewater Monitoring

J. A. B. Coelho, José Gonçalves, and João Paulo Coelho

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Wastewater quality monitoring is essential for protecting public health, preserving ecosystems and ensuring regulatory compliance. Traditional sampling methodologies, based on human intervention, are often insufficient to detect intermittent pollution events. This paper presents the development of a cyber-physical automatic water sampling system designed for deployment at geographically distributed small-scale wastewater discharge points. On top of the mechatronic infrastructure, a digital twin architecture based on FIWARE and NGSI-LD context management is proposed to enable scalable data integration, semantic interoperability, predictive analytics, and pollution tracking. The paper describes the overall system architecture, the data modelling approach and an envisioned digital twin framework that combines embedded sensing with machine learning models. The proposed approach demonstrates how cyber-physical systems and digital twins can enhance wastewater surveillance and environmental governance.

Keywords: Wastewater monitoring · Digital Twin · FIWARE · Water pollution

Evolutionary Hyperparameter Optimization of Neural Networks in Smart Grid Applications

Pedro Jorge Oiiveira Câmara, Paulo Roberto Mann Marques Júnior, and Carolina Gil Marcelino

Federal University of Rio de Janeiro, Brazil

Energy generation and consumption are of global interest, and the United Nations Sustainable Development Goals aim to ensure reliable, sustainable and economically viable access to electricity by 2030. Accurate prediction of electricity consumption is crucial for strategic planning, while the transition to Smart Grids requires continuous assessment of system stability, as grid price fluctuations depend on physical network reliability. Artificial Intelligence models can address these challenges: regression to estimate consumption and classification to evaluate Smart Grid stability. Neural network models are well-established for both tasks due to their capacity to capture complex nonlinear relationships; however, their performance depends critically on carefully chosen hyperparameters. This work applies population-based algorithms to optimize these hyperparameters, comparing them against Optuna, a state-of-the-art Bayesian optimization framework, on publicly available datasets. Results show that metaheuristic methods provide robust performance for regression, with Particle Swarm Optimization achieving the lowest mean RMSE, while for classification both approaches are statistically equivalent, suggesting these methods represent a viable alternative for hyperparameter tuning in energy-related machine learning applications.

Keywords: Artificial Intelligence · Evolutionary algorithms · Energy

Deep learning applied to European canker detection and segmentation

Bruno Raphael Facundo, Pedro João Rodrigues, Pedro Luiz de Paula Filho, and Silvio André Meirelles Alves

Embrapa - Centro Nacional de Pesquisas em Uva e Vinho, Brazil

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Federal University of Technology of Paraná (UTFPR), Brazil

European canker, caused by Neonectria ditissima, is a major phytosanitary disease affecting apple orchards, leading to branch dieback and yield reduction. Accurate lesion detection and segmentation are essential for timely intervention, yet automated analysis remains challenging due to irregular lesion morphology and complex bark textures. This work evaluates deep learning approaches for branch-level European canker detection and segmentation under synthetic orchard conditions. A synthetic dataset was generated by compositing laboratory-acquired branch images onto orchard video backgrounds to simulate field-like visual complexity. YOLOv8 and YOLOv12 models were evaluated for detection and instance segmentation, while Mask2Former and detection-guided SAM2 were assessed for refined lesion delineation. Performance was measured using mAP50 and mAP50-95 for detection, and Intersection over Union (IoU) and Dice coefficient for segmentation. YOLOv8 achieved strong detection performance, reaching mAP50 = 0.9536 in the best configuration, whereas Mask2Former obtained the highest segmentation accuracy (IoU = 0.7514; Dice = 0.8580). Detection-guided SAM2 significantly improved segmentation over single-stage YOLO models but did not surpass fully supervised transformer-based training. Inference analysis revealed a clear trade-off between segmentation accuracy and computational efficiency, with YOLO operating near real-time and transformer-based models incurring higher processing cost. These findings highlight the effectiveness of transformer architectures for irregular lesion delineation while emphasizing practical considerations for real-time deployment in precision agriculture.

Keywords: European canker · Apple tree disease · Object detection · Instance segmentation · Transformer-based segmentation · Precision Agriculture

Assessment of Logistics Performance in European Union Countries

Luis Enrique Krulikowski, and Clara Bento Vaz

Faculdade de Tecnologia de Ourinhos, Brazil

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Logistics performance is a central pillar for economic competitiveness, especially within the integrated context of the European Union (EU). Although the World Bank's Logistics Performance Index (LPI) is the primary benchmarking tool, its aggregation methodology may conceal important nuances in countries' performance. This study conducts a multi-model analytical framework for logistics performance assessment of the 27 EU countries, using the most recent data (LPI 2023), applying the Benefit-of-the-Doubt (BoD) model to evaluate logistics performance and the TOPSIS method to create a multicriteria performance ranking. The analysis reveals that a group of seven countries, including Sweden, the Netherlands, and Germany, reaches the best performance in the BoD model, while nations such as Bulgaria and Cyprus are the worst performers. Additionally, comparing rankings across various methods reveals some dissociation between market perception, as measured by the LPI, and performance calculated from the structure of its sub-indicators. This illustrates that a single assessment model may mask weaknesses and potentialities. To ensure transparency and detailed exploration, all results were consolidated into an interactive web application built with Streamlit library.

Keywords: Logistics Performance Index · Benefit of the Doubt · Data Envelopment Analysis · TOPSIS · European Union

Deep Reinforcement Learning for Intelligent Resource Management in Distributed Cloud Systems

Janio Otoni, and Rui Pedro Lopes

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Traditional resource managers in distributed cloud systems rely on fixed heuristics (Round-Robin, Best-Fit) that struggle with modern workloads. We present a Deep Reinforcement Learning (DRL) framework that shifts resource allocation from static rules to adaptive policies balancing three competing objectives: execution time, energy efficiency, and SLA preservation. Our approach combines recurrent neural networks (LSTM) to handle partial observability with Proximal Policy Optimization (PPO) for stable policy learning. We validate on synthetic workloads derived from Google Cluster Traces v2 across 20 heterogeneous simulated servers. Compared to heuristic baselines (Round-Robin, Best-Fit, Random, Greedy-CPU), our PPO-LSTM agent achieves 4-52% improvements in scheduling reward and 31% energy reduction versus Round-Robin. While hardware constraints limit scale (4GB VRAM) and validation remains simulation-based, the results demonstrate DRL's potential for adaptive resource management with discussion on limitations and future work directions.

Keywords: Deep Reinforcement Learning · Resource Management · Cloud Computing · Task Scheduling · PPO

Hierarchical Temporal Convolution for District Sequencing in Large-Scale Last-Mile Delivery

Farzam Salimi, Ana Rita Simões do Paço E Vaz, and António Galrão Ramos

INESC TEC, ISEP, Portugal

Hierarchical last-mile delivery frameworks decompose city-scale planning into districting, district sequencing, and detailed routing. Within this structure, the sequencing stage strongly constrains downstream routing quality and runtime, yet many practical policies remain static or depend only on first-order transitions. This paper proposes a Hierarchical Temporal Convolutional Network (H-TCN) for multi-resolution district sequencing over a three-level polygon hierarchy. The model uses parent-conditioned dilated causal convolutions, providing a controllable receptive field, parallel training, and a lightweight deployment profile. The experimental protocol is based on a Porto delivery setting in which delivery points are assigned to hierarchical districts, route-level sequences are extracted at Levels 1-3, and the model is evaluated through teacher-forced token prediction, constrained autoregressive rollouts, macro-order distance proxies, and efficiency metrics. The results show that H-TCN is particularly effective at the coarsest hierarchical level, where it improves clearly over strong practical baselines, while remaining competitive at finer levels where the prediction problem becomes more locally constrained. The paper therefore presents temporal convolution as a credible approach for hierarchical district sequencing in real delivery operations.

Keywords: Last-mile delivery · Hierarchical sequencing · Temporal convolution · Vehicle routing · District sequencing

Proposal of a Multidimensional Framework to Analyse Regulation in Artificial Intelligence on Medical Devices

Beatriz Mafra Cabrera, Paula Odete Fernandes, and João Paulo Teixeira

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Applied Management Research Unit (UNIAG), Portugal

This study explores the governance of digital health, examining the growing complexity of public policies and the regulation of artificial intelligence (AI) applied to medical devices, particularly in light of consolidated regulatory frameworks such as Regulation (UE) 2024/1689 (AI Act), Regulation (UE) 2017/745 (MDR), the National Artificial Intelligence Initiative Act in the United States, and Order No. 797 in China. The central problem addressed is how to measure the impact of these regulations on the development, innovation, compliance, and safety of intelligent medical devices while balancing patient protection and technological dynamism. The objective of this study is to propose a multidimensional framework of metrics to assess regulatory effects in an integrated manner across scientific, clinical, social, and economic dimensions. Methodologically, the study adopts a qualitative approach grounded in related works and documentary analysis of international regulatory instruments. As its main result, the study develops a framework structured around four dimensions: innovation and scientific production market access and regulatory compliance clinical and societal impact and stakeholder perception and ecosystem vitality. It concludes that evaluating regulatory impact requires a systemic and continuous approach, providing support for adaptive, evidence-based governance in the application of AI in healthcare.

Keywords: Technology governance · Regulatory compliance · Digital health innovation

Graph-based Tracking Algorithm for Musical Tempo Detection

Luis Morales, Lorenzo J. Tardón, Sergio Guillén, Ana M. Barbancho, and Isabel Barbancho

ATIC Research Group, ETSI Telecomunicación, Universidad de Málaga, Spain

Musical tempo estimation is a fundamental task in Music Information Retrieval (MIR), enabling applications such as beat tracking, rhythm analysis, music synchronization, automatic accompaniment, and interactive multimedia systems. In this work, we present a scheme for the application of a graph-based algorithm for tempo tracking that models tempo evolution as an optimal path in a temporal graph derived from tempogram representations. The resulting tempo trajectory is obtained through a Dijkstra-like shortest-path optimization that identifies the most plausible sequence of tempo candidates across time. The proposed method first extracts a novelty curve from the audio signal and then computes a Fourier-based tempogram to obtain tempo candidates over time. These candidates are organized into a directed graph in which nodes represent tempo hypotheses and edges encode temporal continuity constraints. The final tempo trajectory is obtained using a shortest-path strategy that balances candidate salience and temporal smoothness. The parameters governing candidate extraction and graph costs are optimized using Particle Swarm Optimization (PSO). Experimental evaluation on a small set of representative classical recordings shows that the proposed approach yields reliable tempo trajectories in rhythmically stable signals and remains musically interpretable in the presence of harmonic tempo ambiguity. The results should be understood as a controlled proof of concept rather than as a large-scale benchmark. Beyond tempo tracking, the proposed framework can support applications such as rhythm structure analysis, music synchronization, interactive performance systems, and multimodal studies of rhythmic perception, including future extensions toward EEG-informed analysis of neural entrainment to musical rhythm.

Keywords: Tempo Tracking · Music Information Retrieval · Tempogram · Graph Optimization · Particle Swarm Optimization

A Signal Processing Pipeline for Electromyography: IPB-FosEMG-DB case study

Caio Oliveira, Luiz Eduardo Luiz, and João Paulo Teixeira

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Associate Laboratory for Sustainability and Technology in Mountains Regions (SusTEC), Instituto Politécnico de Bragança, Portugal

Surface electromyography (sEMG) is a pivotal technique for non-invasive analysis of neuromuscular function. However, the inherently stochastic and heterogeneous nature of the signal, coupled with susceptibility to power-line interference, motion artefacts, and inter-subject biological variability, imposes significant challenges on data quality and reproducibility. This paper describes the digital signal processing (DSP) pipeline utilised for the acquisition, conditioning, and adaptive segmentation of forearm sEMG signals during the construction of the IPB-FosEMG-DB dataset. The methodology integrates spectral analysis for precise filtering, full-wave rectification, and linear envelope extraction, followed by a custom adaptive segmentation algorithm that handles variability in voluntary contractions. The pipeline was validated using the dataset, comprising data from 15 healthy volunteers performing isometric and dynamic hand compression tasks. The results demonstrate the pipeline's efficacy in identifying activation periods, providing a reproducible, modular framework for future research leveraging this database.

Keywords: Surface Electromyography · Signal Processing · Adaptive Segmentation · Dataset Generation · Biosignals

Distributed AI Training Platforms: a Comparison

Tiago Cerqueiro, Rui Pedro Lopes, and José Rufino

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Training large-scale artificial intelligence models has become increasingly critical, requiring distributed infrastructures capable of efficiently coordinating multiple devices. This paper presents a comprehensive comparative analysis of three distributed deep learning training platforms: PyTorch Distributed Data Parallel (DDP), Apache Spark, and Determined AI. We evaluate their performance, resource management capabilities, and usability in organizational environments. Experiments on a three-node cluster, using BERT-tiny for sentiment classification on the IMDB dataset, demonstrate that PyTorch DDP achieves the best absolute performance (499 seconds for 20 epochs on 2 GPUs), while Determined AI introduces a 21% overhead but provides superior cluster management, automated scheduling, experiment tracking, and fault tolerance. Apache Spark presents significant overhead (187%) but integrates naturally into existing data processing pipelines. Our findings reveal three critical trade-offs: performance vs. ease of use, control vs. automation, and specialization vs. generality. Framework selection depends critically on context: DDP is ideal for individual researchers prioritizing speed, Determined AI suits shared environments requiring reproducibility and centralized management, and Spark serves scenarios where training is integrated into broader big data workflows.

Keywords: Distributed Training · Deep Learning · Machine Learning · Resource Management · Parallel Computing

A Surrogate-Based Workflow for Indoor Hydroponic Lettuce Optimization

Miguel Afonso Beckers, Pedro Luiz de Paula Filho, Glauco Vieira Miranda, Arnaldo Candido Junior, and Pedro João Rodrigues

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Universidade Estadual Paulista, Brazil

Universidade Tecnológica Federal do Parná, Brazil

Optimizing indoor hydroponic lettuce production is important for enhancing its economic viability. This study applied an Artificial Neural Network (ANN) combined with a brute-force algorithm for optimization. The gain in Fresh Matter Mass (FMM) was set as the optimization goal, while irrigation time off (IT-Off), electrical conductivity (EC), and nutrient-solution potential of hydrogen (pH) were the optimization parameters. The method involved four steps: conducting exploratory cycles to generate a database training an ANN surrogate model using the brute-force algorithm to find the optimal combination in the model and conducting validation plantings to verify the found combinations. The exploratory plantings generated results with variations, which were normalized, and additional lighting data were collected. In general, the combination of 25 minutes of IT-Off, 2.2 mS/cm of EC, and 5.5 of pH was predicted to produce the best average value of FMM, 113.89 g. Two validation plantings were conducted: the first exceeded the prediction by 34.43%, while the second was 17.20% below the expectation. The first result surpassed the exploratory plantings' values, but the second did not. These findings support the feasibility of the proposed surrogate-based workflow, while also indicating that stronger environmental control and additional validation cycles are required to demonstrate robust reproducibility.

Keywords: Modified hydroponic shipping container · Controlled environment agriculture · Artificial neural networks · Brute-force algorithm

Case Study of a Joint Routing and Scheduling Optimization Approach for Multi-Vehicle Job Shops

João Guilherme Martins Silva, João Braun, João Fabro, José Lima, and Ana I. Pereira

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Federal University of Technology of Paraná (UTFPR), Brazil

When industrial warehouse logistics are modeled using job shop/flow shop formulations, it is common for transportation to be simplified to the shortest physical path between origin and destination, implicitly assuming that vehicles operate in a wide enough space that collision cannot occur (or are easily avoidable). This assumption creates a reality gap in congested and tight shop floors, where shared paths and destinations require some form of conflict resolution to ensure collision-free routing, which inevitably modifies paths and causes delays in the logistic chain. This work aims to quantify the impact of this simplification and evaluate whether a tighter integration between scheduling and multi-vehicle routing is capable of bringing significant improvements, using the rules of the RobotAtFactory 4.0 competition as a concrete study environment. A genetic algorithm is used for scheduling optimization, while the Time-Extended A* (TEA*) algorithm is applied for multi-vehicle path planning. After analysing 1458 scenarios, it was determined that the majority of the simplified solutions (70%-86%) would result in a collision in a real scenario, requiring path corrections after optimization. The joint scheduling and routing solution presented statistically inconclusive efficiency gains compared to the simplified solution with path corrections, possibly due to the simplicity of the studied environment.

Keywords: Hybrid Flowshop · Scheduling · Multi-robot Path Planning · Routing · Time-Extended A* · TEA* · Genetic Algorithm · GA

Genetic Programming for Crown Delimitation: An Exploratory Study in Remote Sensing

Raphael Britto, Felipe Gimenez, João Paulo Castro, José Lima, and Ana I. Pereira

Centro de Investigação de Montanha (CIMO), Portugal

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Tree counting plays an important role in forest or agricultural management. In this context, remote sensing methods combined with crown delimitation are a typical approach that facilitates the coverage and separate structural estimation of each plant. Optimization techniques are typically used to deal with canopy segmentation, often involving inputs derived from LiDAR data. Intelligent algorithms related to neural architectures are widely explored, usually ignoring concerns about process explainability. In contrast, optimization strategies grounded on procedural transparency are not much promoted. This work aims to introduce Cartesian Genetic Programming (CGP) in tree counting tasks, addressing a performance comparison with a well-known Fully Convolutional Network (FCN) model. The main differences between the two approaches will be highlighted, focusing on precision by analyzing MCC values and exploiting black- and white-box behaviors. The CGP method, using the Kartezio framework, completely outperformed the FCN's, using an adapted U-Net architecture. Kartezio performed better within the 40 experiments carried out, achieving 0.148 and 0.584 as maximum and minimum MCC values, while U-Net's ranged -0.024 to 0.553. In addition, inputs with two different spatial resolutions were analyzed, framing possibilities for methodology's scalability in future research.

Keywords: LiDAR · CHM · Segmentation · FCN · CGP · MCC

Multi-Step Heart Rate Forecasting using millimetre-Wave Radar and Environmental Context

Rebeca B. Kalbermatter, José Lima, António Valente, and Ana I. Pereira

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

University of Trás-os-Montes and Alto Douro (UTAD), Portugal

Contactless vital sign monitoring via millimetre-Wave (mmWave) radar is an emerging technique for unobtrusive healthcare. Nevertheless, although real-time detection is a popular research topic, forecasting vital signs in multisensory indoor settings remains an unexplored area. This research examines the effects of data engineering choices and multimodal sensor fusion (physiological and environmental) in heart rate (HR) forecasting via Long Short-Term Memory (LSTM) networks. Framed as a time-series regression task, the effects of presence filtering, memory length, and forecasting intervals on a longitudinal dataset spanning four months to capture the seasonal drift of the concept were analysed. The results demonstrate that incorporating periods of subject absence in heterogeneous data acts as an implicit regulator, which enhances the overall robustness of the model. Moreover, the temporal memory requirement for effective HR forecasting quickly saturates at two minutes, optimizing the viability of edge computing. Most importantly, multi-step forecasting results reveal an optimal temporal window in which, while immediate HR dynamics are highly autoregressive and contaminated by environmental noise, multimodal environmental context significantly improves model performance at an intermediate time horizon.

Keywords: Heart Rate Prediction · MmWave Radar · Multimodal Sensor Fusion · Time-Series Prediction · Smart Healthcare · LSTM

Extrinsic Calibration of Heterogeneous 3D Sensors Using Global Registration and Robust Refinement

Vinicius Grilo, João Braun, Ana I. Pereira, and José Lima

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

This work proposes a processing pipeline to estimate the rigid-body transformation between two heterogeneous 3D sensors from point clouds captured in the same scenario and at the same instant. The main objective is the extrinsic calibration of sensors, required for sensor fusion, mapping, and 3D reconstruction. The approach combines a global initialisation, based on FPFH geometric descriptors and RANSAC estimation, with a robust iterative refinement phase using M-estimators, reducing the impact of incorrect matches and outliers. Before registration, preprocessing with distance filtering and voxel grid downsampling is applied to equalise densities between sensors and improve the numerical stability of the alignment. Validation is conducted in two scenarios: (i) a synthetic experiment with known transformation, allowing quantitative error assessment, and (ii) a real scenario with a Unitree GO2 robot equipped with Robosense RS-16 LiDAR and Intel RealSense D435i RGB-D camera. The results indicate that the combination of global initialisation and robust refinement increases the stability of the process and improves the quality of the alignment in realistic conditions with density differences, occlusions, field of view variation, and the presence of outliers.

Keywords: Extrinsic calibration · Point cloud registration · Robust optimization

Dynamic Difficulty Adjustment in Virtual Reality Rehabilitation Game Through Reinforcement Learning

Isaac Van-Deste, Júlio Castro Lopes, and Rui Pedro Lopes

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Schizophrenia is a multifaceted complex mental disorder, that often leads to loss of physical and cognitive skills, reducing social interaction, with a high relapse rates, that turns rehabilitation a long-term challenge. Virtual Reality has been emerging as a promising tool for rehabilitation, since it provides a safe, controlled, and motivating therapeutic scenarios. Upon this, the present paper introduces a VR-based serious game, capable of dynamically adapting its difficulty in response to physiological and behavioural signals of the patient. To achieve this solution, several different data from the patient, such as heart rate, body motion rate, and facial expression recognition were continuously monitored and processed to infer the patient's stress levels in real time. The stress estimate is then integrated into a Dynamic Difficulty Adjustment mechanism, formulated in a Reinforcement Learning algorithm. A Deep Q-Network agent was trained within a simulated environment to learn optimal policies that balance challenge and engagement, ensuring that rehabilitation sessions remain neither frustrating nor trivial. Validation of the agent was done, by causing interference into the agent, with gRPC. Both training and validation of the agent confirms its ability to response accordingly to stress peaks by lowering difficulty, demonstrating the systems robustness and real-time adaptability.

Keywords: Artificial Intelligence · Reinforcement Learning · Dynamic Difficulty Adjustment · Virtual Reality

Machine Learning for Diabetes Mellitus Prediction: A Comparative Analysis of Ensemble and Single Models

Gabriel A. Leite, Beatriz Flamia Azevedo, Maria F. Pacheco, Florbela P. Fernandes, and Ana I. Pereira

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Polytechnic Institute of Bragança, Portugal

Diabetes Mellitus is a rapidly growing disease that presents numerous risks and can be fatal in severe cases. This study proposes an approach that combines and compares several well-known machine learning models - including Random Forest, Gradient Boosting, Multi-Layer Perceptron, Support Vector Machine, and ensemble methods (Bagging, Stacking, Boosting, and Voting) - to analyze patient data in the context of this disease. The results showed that boosting-based models, such as CatBoost, XGBoost, and Gradient Boosting, achieved the highest performance, with Area Under the Curve values of 0.8532, 0.8531, and 0.8529, respectively. Finally, a voting ensemble was used to combine the strengths of Random Forest, Gradient Boosting, and MLP to identify the most relevant patient features. The top three predictors for determining whether a patient is affected by the disease were High blood pressure, Health status, and Ever Told Blood Pressure High, with respective importance values of 12.10%, 11.28%, and 7.20%.

Keywords: Diabetes Mellitus · Boosting · Random Forest · Medical · Classifier Combination

Enhancing Small Object Detection in UAV-Based Olive Monitoring Using YOLO12 and SAHI

Maurício Morais, João Mendes, José Lima, and Ana I. Pereira

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

This work presents a UAV-based workflow for detecting olives directly on tree canopies using lightweight YOLO12 detectors combined with Slicing Aided Hyper Inference (SAHI). High-resolution lateral video of individual trees is converted into tree-level crops by a YOLO12 tree detector, which attains a peak F1 Score of 0.99 and mAP50 of 0.990 on the validation split. Three YOLO12 variants (nano, small, and medium) are trained on a custom dataset of 1033 tree-level crops and 13.911 annotated olives, and evaluated under conventional and SAHI-based inference. On the fruit-detection task, YOLO12s with SAHI achieves the best results, with an mAP50 of 0.907 and F1 of 0.863, outperforming all conventional variants, particularly on small and clustered fruits. Runtime experiments show that SAHI adds moderate latency relative to 640-pixel inference but preserves VRAM usage while avoiding the accuracy loss observed with upscaling to 1280 or 1920 pixels, making the proposed workflow a practical basis for per-tree olive yield estimation.

Keywords: Olive detection · Small-object detection · YOLO12 · Slicing Aided Hyper Inference (SAHI) · Precision Agriculture

Exploring the Gap Between Local Sensors and Open-Source Weather Data

Bip Diwakar, João Mendes, Inês Sena, Hugo Cardoso, Florbela P. Fernandes, Ana I. Pereira

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Skillent, Coimbra

Precision viticulture depends on dense sensor networks to monitor environmental variables that influence vine physiology and irrigation management. However, the installation and maintenance of physical sensors involve considerable cost and operational effort. This study evaluates whether freely available online meteorological data can act as reliable virtual sensors under strict validation conditions, to provide critical agronomic data when in-field vineyard sensors are unavailable or fail. Five external data sources were compared, and Open-Meteo was selected based on correlation strength and bias stability. Fourteen regression models were tested across eleven vineyard parameters using time-aware validation on 5.016 hourly observations. Results show that atmospheric pressure, soil temperature at multiple depths, UV radiation, and leaf temperature can be reliably predicted using external data, achieving R^2 values greater than 0.8. In contrast, leaf humidity, soil moisture, and precipitation remain poorly predictable. The findings provide evidence-based guidance for designing cost-effective vineyard monitoring systems.

Keywords: Machine Learning · Precision Viticulture · Virtual Sensing · Time Series Prediction

Multi-View Terrain Traversability Mapping from UAV RGB Imagery for UGV Navigation

João Mendes, Marek Nowakowski, Guido S. Berger, Maurício Morais, José Lima, and Ana I. Pereira

Military Institute of Armoured and Automotive Technology, Poland

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Terrain traversability assessment is a key requirement for safe Unmanned Ground Vehicle (UGV) navigation in unstructured out-door environments. This work presents a passive vision-based framework for traversability mapping using RGB imagery acquired by low-cost Un-manned Aerial Vehicles (UAVs). The proposed approach combines se-mantic segmentation, RGB mosaicking, and probabilistic multi-view fu-sion to generate spatially consistent traversability maps over extended areas. Frame-level segmentation is first performed using a DeepLabV3+ model, and two inference strategies are evaluated: baseline resizing and tiled inference. The predicted navigability probabilities are then pro-jected into a common mosaic reference frame and fused across overlap-ping observations, while a weighted variance map is estimated to char-acterize inter-frame disagreement. Finally, the fused representation is converted into a vehicle-aware passability map, enabling route feasibil-ity analysis between user-defined start and goal locations. Experimental results show that tiled inference improves frame-level segmentation per-formance compared with the baseline approach, while multi-view fusion produces coherent traversability maps that preserve the navigable road structure and support practical path planning. The results suggest that passive UAV RGB imagery can provide an effective and lightweight ba-sis for terrain assessment and navigation support in complex outdoor environments.

Keywords: Traversability mapping · UAV RGB imagery · Semantic segmentation · Multi-view fusion · UGV navigation

Lexical and Discursive Changes in the Era of Generative AI

Inês Sena, and Ana I. Pereira

Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politécnico de Bragança, Portugal

Generative Artificial Intelligence (GenAI) has brought about significant transformations in scientific research and technological development. In addition to its practical applications, these technologies can also influence how scientific knowledge is produced and communicated. In this context, the present study analyzes the lexical and discursive changes in Artificial Intelligence (AI) research associated with the emergence of GenAI. To this end, two distinct time horizons are compared: a pre-GenAI period (2015-2019) and a post-GenAI period (2022-2025). The analysis is based on a corpus of scientific publications indexed in the Web of Science (WoS) database, from which 25.000 papers were selected for each time period. Using different data analysis techniques to identify new terms, emerging words, declining terminology, and patterns of discursive change. The results reveal that although the core vocabulary of AI research remains relatively stable, the post-GenAI period has seen the emergence of new terms directly associated with generative technologies, including concepts related to Large Language Models and conversational systems. In parallel, there is a significant growth in terminology associated with governance, ethics, trust, and the societal impact of AI. Discourse analysis also reveals a shift in the scientific narrative, with greater use of language related to technological transformation, practical implementation, and the need for institutional adaptation. To conclude, the results obtained suggest that the emergence of GenAI is associated not only with the introduction of new technologies but also with observable transformations in the vocabulary and discourse of AI research.

Keywords: Artificial Intelligence · Generative AI · Data analysis · Lexical Changes.

Training Machine Learning Models on Encrypted Data: A Privacy-Preserving Framework using Homomorphic Encryption

Alexandre Marques, Beatriz Sá, Rui Botelho, and Pedro Pinto

ISEP, School of Engineering, Polytechnic of Porto, Porto, Portugal

The use of Machine Learning (ML) for data-driven decision-making often relies on access to sensitive datasets, which introduces privacy challenges. Traditional encryption methods protect data at rest or in transit but fail to secure it during processing, exposing it to unauthorized access. Homomorphic encryption emerges as a transformative solution, enabling computations on encrypted data without decryption, thus preserving confidentiality throughout the ML pipeline. This paper addresses the challenge of training ML models on encrypted data while maintaining accuracy and efficiency by proposing a privacy-preserving framework that leverages Cheon-Kim-Kim-Song (CKKS) for approximate real-number arithmetic. This demonstrates the feasibility of training K-Nearest Neighbors (KNN) and linear regression models on encrypted datasets, ensuring that data remains confidential even during computation. Additionally, an implementation of Multilayer Perceptron (MLP) was tested on the same dataset to evaluate the applicability of the framework to basic neural network models. Experimental results show that models trained under Homomorphic encryption achieve performance metrics comparable to plaintext-trained models, validating the approach. However, challenges such as computational overhead, noise management, and limited support for non-polynomial operations persist. This work lays the groundwork for broader adoption of privacy-preserving ML in real-world applications, balancing security with computational feasibility.

Keywords: Homomorphic encryption · Privacy-preserving Machine Learning · Secure data analytics · Encrypted model training

Authors Index

Keywords Index

Sponsors

Instituto Politécnico de Bragança is a Portuguese University of Applied Sciences with 10500 students and 500 teachers and researchers. It has an internationalization programme with widely recognized success, including the annual mobility of more than 1000 incoming and outgoing students and 300 teaching staff members. IPB is also recognized by its performance in applied research and regional engagement.

CeDRI stands for "Research Centre in Digitalization and Intelligent Robotics" and its mission is to contribute to the development, application and transfer of scientific knowledge related to robotics, intelligent systems, and information and communication technologies, in the scope of the digitalization and automation of industrial systems, strengthening the scientific and technological system.

University of Málaga (UMA) is a public university in Andalusia, Spain, founded in 1972. It offers a wide range of undergraduate, master's, and doctoral programs across sciences, engineering, humanities, health, and social sciences. Located in the city of Málaga, it is known for its strong research activity, international collaborations, and connection with industry. UMA plays an important role in regional development and innovation, especially in technology and applied sciences.

STEP Project aims to drive excellence and enhance the research capacity and reputation of Instituto Politécnico de Bragança (IPB) in the widening country of Portugal, by excelling in the domain of STEM (Science, Technology, Engineering and Mathematics) and EDI (Equality, Diversity and Inclusion) values through cross-fertilisation.

Fundação para a Ciência e a Tecnologia, is the Portuguese public agency that supports science, technology and innovation, in all scientific domains, under responsibility of the Ministry for Science, Technology and Higher Education.

Springer is a leading global scientific, technical and medical portfolio, providing researchers in academia, scientific institutions and corporate R&D departments with quality content through innovative information, products and services.