CIABIOMED 2025

ETSII – US

OCTOBER 23-24, 2025

https://2025.iabiomed.org

BOOK OF ABSTRACTS

BIOMED
Sociedad Española de
Inteligencia Artificial en Biomedicina

Event sponsors

Sponsors of IABiomed

Event organizers and collaborators

BIOMED
Sociedad Española de Inteligencia Artificial en Biomedicina

Index

P1-01

IA aplicada a la pangenómica para explorar la constante guerra entre virus y bacterias

Antonio J. Pérez Pulido
Andalusian Center for Developmental Biology (CABD) and Pablo de Olavide University.

Bacteriophages (bacterial viruses or phages) exist in a ratio of 1 to 10 with respect to bacteria, and it is estimated that they eliminate up to half of these bacteria daily in different natural environments. Bacteria, in turn, defend themselves using different biological weapons, of which there are currently more than a hundred known types.

Our group constructs and analyses bacterial pangenomes using thousands of public genomes of clinically relevant bacteria. These pangenomes are translated into binary matrices of gene presence-absence, with tens of thousands for each, to which information about phages, their receptors (proteins they use to enter the bacterium) and the defence and anti-defence systems of both are added. Finally, we apply machine learning techniques such as random forest with XGBoost and CN2 rule induction as tree-based algorithms, along with other standard data mining methods, to identify marker genes related to the characteristics of interest.

In this way, we have been able to establish relationships between bacterial adaptive immunity defence systems, CRISPR-Cas, and phages of different species, as well as find specific defence systems of groups of bacteria of great clinical concern. These results are useful for better understanding the genomic plasticity of this bacterium and could be used in phage therapy, the alternative to antibiotics that is becoming increasingly important.

P2-01

Si investigas como en 2010, la IA escribirá tu obituario científico

Juan Mora Delgado
Facultativo Especialista de Área de Medicina Interna. UGC de Medicina Interna y Cuidados Paliativos. Hospital Universitario de Jerez de la Frontera.

Artificial intelligence is not here to replace researchers, but to free them from what takes up most of their time: endless searches, statistical tables, methodological reviews, and impossible journal formats. In this experience, we will take a practical tour of AI co-pilots created specifically for biomedical research: from generating research questions and advanced bibliographic searches to calculating sample size, scientific writing, and preparing articles for high-impact journals.

The goal is to show, with real examples requested by the live audience, how any healthcare or biomedical professional can shorten years of research work and focus on the essentials: science and patients.

I1-01

AI at MSD: vision and strategy from the biopharmaceutical sector.

Helena González, Javier Dúctor
MSD Spain.

In this session, we will review how artificial intelligence is transforming research and development in the pharmaceutical industry, from drug discovery to clinical trials and patient access. We will share representative use cases, identify challenges and opportunities, and propose avenues for collaboration to accelerate the impact of AI in biomedicine.

I2-01

Onesait Healthcare CLAI: Health Co-pilot for doctors and patients: RAG and GraphRAG on interoperable history, help sources, and living science.

Juan Manuel Martínez Pérez, Andrés Carretero Sosa
MINSAIT

Conversational co-pilot designed to assist doctors and patients in their daily tasks using multi-channel access. Using a RAG and GraphRAG approach, the system combines semantic retrieval (graph) and vector retrieval (notes, guides, SOPs), orchestrated by agents that execute clinical tools (FHIR queries on medical records, bibliographic search, smart forms) and return responses with verifiable citations, explainability, and quantification of uncertainty.

SA1-01

An Iterative Random Forest Framework for Statistical Feature Selection in High-Dimensional Biomedical Data: A Case Study on Alzheimer's Diagnosis

Pablo Zubasti[1], Miguel A. Patricio[1], Antonio Berlanga[1], and José M. Molina[1]
[1]*Universidad Carlos III de Madrid, ROR: https://ror.org/03ths8210. Computer Science and Engineering Department. Applied Artificial Intelligence Group, Spain*

This study presents a robust methodology to address the challenge of high-dimensionality in clinical datasets, a pervasive issue in biomedical informatics. Leveraging the predictive capabilities of Machine Learning—particularly in healthcare applications—has shown considerable promise. However, the presence of redundant or low-relevance features in training data often hampers model efficiency and inflates both computational and clinical costs. These concerns are especially critical in medical scenarios, where data acquisition may involve expensive or invasive procedures. To mitigate these challenges, we introduce an iterative framework that employs multiple Random Forest models and a Gaussian modeling approach to statistically assess and rank attribute importance. Through repeated training cycles, our approach identifies non-contributory and statistically insignificant features, enabling their systematic exclusion.

SA1-02

Towards an Explainability Agent: Leveraging LLMs to Interpret LIME Outputs

Belén Vega-Márquez[1], Cristina Rubio-Escudero[1], and Beatriz Pontes-Balanza[1]
[1]*Department of Computer Languages and Systems, University of Seville, 41012 Seville, Spain*

In clinical decision-making, the adoption of machine learning models requires not only high predictive performance but also transparent, trustworthy explanations. This work presents a hybrid explainability framework that combines Local Interpretable Model-agnostic Explanations (LIME) with a Large Language Model (LLM) to generate natural language explanations of classification outputs in healthcare applications. While LIME provides local feature attributions for individual predictions, these are often difficult to interpret by non-technical clinical staff. Our system leverages an LLM to translate LIME outputs into fluent, domain-aware explanations that align with the reasoning needs of healthcare professionals. We evaluate the method on two medical datasets involving patient risk classification tasks and obtain promising results in terms of interpretability and consistency with the model's decision logic. However, the current evaluation lacks validation by external healthcare professionals, which we identify as an essential next step. Despite this, our approach represents a strong foundation for the development of adaptive explainability agents in clinical contexts. We discuss its potential impact on future decision support systems and propose directions for evolving toward interactive, trustworthy, and user-aligned AI tools in medicine.

SA1-03

Uncovering cardiac risk patterns: visualization and interpretation via probabilistic topographic mapping

Martha Ivon Cardenas[1], Pedro Jesús Copado[1], and Caroline König[1]
[1]*Soft Computing Research Group (SOCO) at Intelligent Data Science and Artificial Intelligence (IDEAI-UPC) Research Centre, Universitat Politècnica de Catalunya (UPC Barcelona Tech), Jordi Girona 1-3, 08034 Barcelona, Spain*

Cardiovascular disease (CVD) remains a leading cause of mortality, yet accurate and interpretable risk prediction poses significant clinical challenges. While machine learning (ML) models show promise in identifying complex patterns, their adoption in healthcare is often limited by a lack of transparency. This paper proposes a novel explainability framework integrating Generative Topographic Mapping (GTM) and Kernel Generative Topographic Mapping (kGTM) to visualize and analyze latent cardiac risk patterns. Unlike traditional dimensionality reduction methods, our approach unifies probabilistic modeling, visual explainability, and clinically grounded subgroup analysis within a single interpretable latent representation. We employ a structured dataset of 1,000 patients with 12 clinical features to demonstrate how GTM and kGTM project high-dimensional patient data into a structured 2D space. This methodology enables direct visualization of patient clusters, classification boundaries, and the influence of individual features, thereby uncovering interpretable regions and clinically meaningful feature interactions directly from the topographic map. This approach provides robust and trustworthy perspectives on ML model behavior for cardiovascular risk assessment. Our dual use of GTM and kGTM allows for adaptive modeling, facilitating the identification of outliers and borderline cases. This dual nature enhances understanding of both the feature space distribution and the uncertainty of the model's decisions, enabling partial explainability by highlighting relevant features and structural patterns in the data.

SA1-04

Explainable Deep Learning Techniques for Medical Image Analysis: A Systematic Review of Diabetic Foot Ulcers, Breast Cancer, and COVID-19

Zinah Mohsin Arkah[1,2], Beatriz Pontes[1], and Cristina Rubio[1]
[1]University of Seville, Seville, Spain, [2]University of Information Technology and Communications, Baghdad, Iraq

Advances in deep learning (DL) have transformed the analysis of medical images for conditions such as diabetic foot ulcers (DFU), breast cancer, and COVID-19. Although these models are highly accurate in their predictions, their limited transparency hinders their clinical use. Described artificial intelligence (XAI) methods, such as Grad-CAM, SHAP, and LIME, have been created to improve interpretability and foster greater trust in these models. This systematic review critically examines recent studies that combine deep learning and explainable AI to address three major healthcare challenges: detecting diabetic foot ulcers, breast cancer imaging, and COVID-19 diagnosis using various imaging techniques, including DFU images, mammography, computed tomography, and Xrays. These studies showcase a range of convolutional neural network (CNN) models, from lightweight, custom-designed models (such as LW-CORONet and COVID-XNet) to more advanced models, including ResNet, DenseNet, and Inception. Our findings suggest that interpretable models can achieve accuracies above 95% while also providing clinicians with both visual and quantitative insights. Challenges include limited dataset diversity, insufficient external validation, and the lack of standardized metrics for interpretability. Overcoming these issues is crucial for encouraging the use of AI tools in clinical practice. The review concludes with recommendations for future research aimed at enhancing model robustness, integrating multimodal data, and enhancing clinical utility. These methods, including Grad-CAM, SHAP, and LIME, are crucial for enhancing the interpretability of deep learning models, enabling clinicians to understand how predictions are made. By increasing transparency and trust in AI systems, these techniques can ultimately improve patient outcomes and facilitate a smoother integration of technology in healthcare.

SB1-01

Fairness-Aware Machine Learning for Biomedical Prediction: Evaluating and Correcting Bias in Gallstone Diagnosis Models

Caroline König[1], Martha Ivon Cardenas[1], Pedro Jesús Copado[1], and Alfredo Vellido[1]
[1]*Soft Computing Research Group (SOCO) at Intelligent Data Science and Artificial Intelligence (IDEAI-UPC) Research Centre, Universitat Politècnica de Catalunya (UPC Barcelona Tech), Jordi Girona 1-3, 08034 Barcelona, Spain*

This study presents a fairness-aware machine learning framework for biomedical prediction, illustrated through gallstone disease detection. Using a structured clinical dataset of 320 patients, predictive models, including multilayer perceptron and random forest, are developed via stratified 5-fold cross-validation. Model performance is evaluated using AUC, F1-score, precision, and recall, with subgroup fairness assessed through demographic parity, disparate impact, predictive parity difference, equal opportunity, and equalized odds. Bias linked to age and gender is quantified and corrected using threshold optimization and F1-based recalibration. Postprocessing interventions demonstrate improved fairness metrics while preserving diagnostic accuracy. Our methodology aligns with ethical guidelines and technical expectations under the EU AI Act, emphasizing equitable model behavior in clinical decision-making. The proposed framework contributes to responsible AI deployment in healthcare, supporting inclusion and transparency in high-risk biomedical applications.

SB1-02

An Agentic Architecture for Scalable and Reproducible Data Standardization to OMOP CDM using Declarative Modeling

Alberto Labarga[1]
[1]*Universidad Pública de Navarra, Spain.*

The secondary use of clinical data for biomedical research is a cornerstone of modern medicine, yet its potential is severely constrained by the persistent challenge of data heterogeneity. While the Observational Medical Outcomes Partnership Common Data Model (OMOP-CDM) provides a robust standard for data harmonization, the process of transforming source data into this model remains a significant bottleneck, characterized by manual, resourceintensive, and brittle Extract-Transform-Load (ETL) pipelines. This paper proposes a novel agentic architecture that reframes data standardization from an imperative scripting task to a declarative, AI-augmented knowledge-generation process. Our framework decomposes the complex mapping workflow into a series of discrete tasks executed by specialized, autonomous software agents. This modular architecture orchestrates a powerful synergy between declarative data modeling with the Linked Data Modeling Language (LinkML), reliable Large Language Model (LLM) interfacing via the Boundary AI Markup Language (BAML), and the cognitive capabilities of LLMs for model alignment. The system's primary output is not merely transformed data, but a human-readable and machine-executable mapping specification using linkml-map. This approach establishes a new paradigm for creating scalable, reproducible, and transparent data standardization pipelines, positioning the data engineer as an expert validator of AI-generated knowledge rather than a manual coder of transformation logic.

SB1-03

The Assessment of AI-Based Digital Health Technologies From The Perspective of HTA Bodies. The Case Of AQuAS' AI Assessment Guide.: AQuAS' AI Assessment Guide

Carolina Moltó Puigmartí[1], Susanna Aussó Trias[2], Maria Bretones Vallejo[2], Didier Domínguez Herrera[3], and Rosa Maria Vivanco-Hidalgo[1]
[1]Agency for Health Quality and Assessment of Catalonia (AQuAS), Barcelona, Spain, [2]TIC Salut Social Foundation, Barcelona (FTSS), Spain, [3]Universitat Oberta de Catalunya (UOC), Barcelona, Spain (note: this work was conducted while affiliated with the FTSS).

This paper presents the AQuAS AI Assessment Guide, a methodological framework developed to support the evaluation of digital health technologies (DHTs) that incorporate artificial intelligence (AI). Designed collaboratively by the Agency for Health Quality and Assessment of Catalonia (AQuAS) and the TIC Salut Social Foundation, the guide defines 13 assessment domains encompassing clinical relevance, technical performance, ethical considerations, regulatory compliance, and broader system and societal impact. The guide aims to foster robust, evidence-based and context-sensitive evaluation of AI solutions across different stages of development and diverse use cases. This article outlines the background, development methodology, and structural composition of the guide, and discusses its potential uses and contributions to responsible innovation in health systems.

SB1-04

Enhancing Privacy and Interoperability in Biomedical Research: A SOLID-Based Architecture with LLM Integration

Hugo Lebredo[1], Jorge Álvarez-Fidalgo[1], Rubén del Rey Álvarez[1], and Jose Emilio Labra-Gayo[1].
[1]*University of Oviedo, Department of Computer Science, Oviedo 33007, Spain*

Ensuring patient privacy in the use of personal data within biomedical research datasets remains one of the foremost challenges in the field. The collection of clinical test results by medical institutions poses significant risks, particularly regarding the potential re-identification of individuals from such data. In this context, this paper proposes a twotier architecture with two key functionalities. First, it enables medical institutions to access clinical data stored in patient-controlled personal repositories in external domains, contingent upon patient consent. Second, it equips healthcare providers with the necessary tools to integrate such data into their systems, facilitating the creation of larger and more complex datasets. To implement this framework, we employ semantic technologies such as Shape Expressions (ShEx), alongside widely adopted medical standards and vocabularies, including FHIR (Fast Healthcare Interoperability Resources) and SNOMED CT. These components are integrated following SOLID (Social Linked Data) principles and made accessible via an integrated large language model (LLM) layer. This approach supports the development of a decentralized, secure, and accessible architecture designed for use by healthcare professionals.

SA2-01

Application of Machine Learning Techniques to the Prediction of Hospital Mortality: Beyond Conventional Clinical Models

Pedro López Ruz[1], Belén Vega-Márquez[1], and Beatriz Pontes-Balanza[1]
[1] *Department of Computer Languages and Systems, University of Seville, 41012 Seville, Spain*

This study focuses on comprehensive analysis of cohorts of patients admitted to Intensive Care Units (ICU), with the purpose of evaluating, comparing, and optimizing the metrics used in the prediction of mortality in critically ill patients. Recognizing the inherent complexity of clinical prognosis in these settings, widely used traditional tools such as the APACHE, SOFA, and SAPS scoring systems are examined, identifying their main strengths and limitations. The integration of advanced Machine Learning and Data Science techniques is also proposed in order to develop more accurate and robust predictive models. The methodology contemplates the application of machine learning algorithms, variable selection procedures, cross-validation and performance evaluation through metrics such as AUC-ROC, accuracy, sensitivity, and specificity. The models will be trained using real data from patients in the ICU, ensuring compliance with ethical and confidentiality principles. The model that demonstrates the best performance will be compared with traditional tools, with the aim of determining its ability to improve mortality prediction and support clinical decision making. In addition to improving predictive accuracy, the aim is to ensure the feasibility and applicability of these models in real clinical contexts. In general, this research aims to contribute to the development of innovative technological tools that favor more personalized, efficient, and evidence-based medical care in highly complex hospital units, promoting optimized management of resources, healthcare personnel, and clinical protocols in the ICU.

SA2-02

Evaluation and prediction of the musculoskeletal risks in microsurgery

Daniel Caballero[1], Manuel J. Pérez-Salazar[1], Juan A. Sánchez-Margallo[1], Laura C. Pires-Louça[2], and Francisco M. Sánchez-Margallo[3]
[1]Bioengineering and Health Technologies Unit, Jesús Usón Minimally Invasive Surgery Center, N-521 Road Km. 41.8, Cáceres 10071 Spain, [2]Microsurgery Unit, Jesús Usón Minimally Invasive Surgery Center, N-521 Road Km. 41.8, Cáceres 10071 Spain, [3]Scientific Direction, Jesús Usón Minimally Invasive Surgery Center, N-521 Road Km. 41.8, Cáceres 10071 Spain

Microsurgery enables precise manipulation of small and complex anatomical structures. However, it is a physically and mentally demanding surgical technique. Advances in wearable technology and artificial intelligence (AI) have improved the development of innovative solutions in this field. The objective of this study is to evaluate the surgeon's physiological and ergonomic parameters by using EMG wearable technology during performance in conventional and robotic-assisted microsurgery (RAM). Additionally, this study seeks to generate predictive models for prediction of musculoskeletal risks, such as localized muscle fatigue, during microsurgery. The data were collected over twenty microsurgical sessions, ten were conventional and the remaining ten were RAM performed by four microsurgeons. The data were recorded in five bilateral muscle groups. The data was recorded for the evaluation of the localized muscle fatigue and the development of predictive models. From these data, a general dataset (508,328 records) was generated, for conventional and RAM. Two different preprocessing techniques (scaled and scaled and normalized) were applied. In this dataset, 80% of the data for training and 20% for testing purposes. This work contributes to the better understanding of ergonomic risks in microsurgery, presenting a significant advance in the integration of wearable technology and the implementation of predictive models of musculoskeletal risks in microsurgery. These results demonstrate the goodness and accuracy of the predictive models and they are the starting point to reaching an exhaustive knowledge of the ergonomic risks of microsurgery.

SA2-03

Implementation of an AI-Assisted Telemedicine System in Nursing Homes: Protocol and Preliminary Results

José Luis Ávila-Jiménez[1], Nuria Luque Reigal[2], Manuel Rich-Ruiz[2], Vanesa Cantón-Habas[2], and Sebastián Ventura[3]
[1]Department of Computer and Electronic Engineering, Universidad de Córdoba, Córdoba, Spain, [2]Department of Nursing, Pharmacology and Physiotherapy, Universidad de Córdoba, Córdoba, Spain, [3]Department of Computing and Artificial Intelligence, Universidad de Córdoba, Córdoba, Spain

Population ageing is increasing the demand for complex care in nursing homes, where continuous medical attention is often limited due to the lack of on-site healthcare professionals. The primary objective of this study is to evaluate the feasibility of applying artificial intelligence to telemedicine in residential care, with a particular focus on developing and validating a predictive model for hospital referral. We conducted a multicentre observational study using a telemedicine system equipped with advanced diagnostic devices to manage acute events in 70 nursing homes across Spain. Clinical and socio-demographic data obtained during telemedicine consultations from 5,192 anonymized patients were used to train a deep learning model. Preliminary findings show that the neural network achieved an accuracy of 0.93 and an AUC of 0.90 in predicting hospital referrals, substantially outperforming traditional classifiers. These results provide early evidence of the potential of leveraging data generated through telemedicine for the development of AI-based predictive models in geriatric care. Future work will focus on expanding the dataset, refining the modelling pipeline, and integrating the predictive model into the telemedicine platform to support real-time clinical decision-making in nursing homes

SA2-04

Improving Liver Graft Decision-Making Through AI: Validation with Internal and National Datasets

Miguel Cuende, Beatriz Pontes, Juan M. Castillo Tuñón, Daniel Mateos García, Luis M. Marín Gómez, José C. Riquelme Santos, and Gloria de la Rosa Rodríguez
[1]HPB and Liver Trasplant Unit, Toledo University Hospital, Toledo, Spain, [2]Department of Computer Languages and Systems, University of Seville, Spain, [3]HPB Surgery and Liver Transplant Unit, Virgen del Rocío University Hospital, Seville, Spain, [4]Spanish Liver Transplant Registry (RETH), National Transplant Organization, Spain

Liver graft assessment is a critical and complex step in the transplantation process, traditionally reliant on the subjective judgment of experienced surgeons, based on initial donor data and macroscopic evaluation. In a previous study, we developed a machine learning-based expert system trained on clinical data to assist in the early decision-making process of liver graft suitability. Using donor information collected in the official Liver Donation Protocol (LDP), the model demonstrated promising predictive performance, particularly in identifying transplantable grafts that were otherwise discarded. This study aims to externally validate and expand the model using newly collected internal data and a large-scale external dataset provided by the Spanish National Transplant Organization (ONT). By evaluating the model's generalization to these additional datasets and retraining it with combined data sources, we assess its robustness, adaptability, and potential for broader clinical application. Furthermore, we explore alternative modeling strategies to enhance predictive performance and reliability. This work represents a critical step toward refining and scaling an AI-driven tool to support transplant surgeons in early-stage graft evaluation using standardized donor variables.

SA2-05

Context-Aware AI Agents for Clinical Dialogue Assistance through Large Language Models

Rodrigo Naranjo Pozas[1], Pablo Doblado Mendoza[1], and Belén Vega Márquez[1]
[1] *Department of Computer Languages and Systems, University of Seville, 41012 Seville, Spain.*

This research investigates the efficacy of Artificial Intelligence Agents in processing and responding to personalized, private contextual information. We studied how to implement a system designed to augment an open-source Large Language Model (LLM), such as Llama, Claude, and Gemma, with domain-specific knowledge bases. This augmentation is intended to facilitate the generation of contextually coherent and accurate responses to user queries. The system was developed and tested using different versions of a clinical conversation transcripts dataset between patients and medical professionals, enabling specialized knowledge integration. The architectural framework, built upon LangChain, FAISS, Ollama, and Gradio, demonstrates a simple, modular, scalable, and extensible design. This work helped us to take our first steps into the development of robust AI agents capable of leveraging external knowledge for enhanced conversational intelligence in specialized domains.

SB2-01

Analysis of Gamma Rhythm in the Detection of Photoparoxysmal Responses in Photosensitive Patients

Fernando Moncada Martins[1,3], Víctor M. González, Antonio Gil-Nagel[2,3], Adrián Valls Carbó[4,5], María Antonia Gutiérrez[6], and Pablo Calvo Calleja[6]
[1]Computer Science Department, University of Oviedo, Spain, [2]Electrical Engineering Department, University of Oviedo, Spain, [3]Biomedical Engineering Center, University of Oviedo, Spain, [4]Neurology Service, Ruber International Hospital, Spain, [5]Fundación Iniciativa para las Neurociencias (INCE), Spain, [6]Neurophysiology Service, Cabueñes University Hospital, Spain

The accurate identifiction of Photoparoxysmal Responses (PPR) is clinically important, as these epileptiform discharges are the key biomarkers used in diagnosing photosensitivity. Traditionally, PPR are identified in electroencephalography (EEG) during Intermittent Photic Stimulation, but manual detection is time-consuming and subjective, highlighting the need for reliable automated approaches. Recent studies have shown that Gamma and High-Frequency Oscillations (HFO) are promising biomarkers in epilepsy research, potentially improving the precision of diagnosis. Building on this, our study incorporates higherfrequency analysis into photosensitivity to investigate how it contributes to identifying PPR. We use an unsupervised anomaly detection method based on a Variational Autoencoder trained on EEG recordings from healthy individuals. Results show that excluding Gamma and High Gamma rhythms greatly decreases the model's detection Accuracy from 83% in prior full-spectrum analysis to 63%, supporting the idea that higher frequencies are key in the distinction of PPR, which aligns with broader epilepsy research, where high-frequency activity has proven to be a key marker of pathological brain dynamics. This approach could help integrate high-frequency biomarkers into clinical decision-support systems as well as support automated EEG analysis in other epilepsy syndromes beyond photosensitivity.

SB2-02

Application of machine learning and deep learning methods on ECG sensor data to predict stress levels in minimally invasive surgery

Daniel Caballero[1], Manuel J. Pérez-Salazar[1], Juan A. Sánchez-Margallo[1], Ismael Diaz-Romero[2], and Francisco M. Sánchez-Margallo[3]
[1]Bioengineering and Health Technologies Unit, Jesús Usón Minimally Invasive Surgery Center, N-521 Road Km. 41.8, Cáceres 10071 Spain, [2]School of Technology, University of Extremadura, Av/ de la Universidad s/n, Cáceres 10003 Spain, [3]Scientific Direction, Jesús Usón Minimally Invasive Surgery Center, N-521 Road Km. 41.8, Cáceres 10071 Spain

This study aims to predict the values of ergonomic (body position, angular rate), kinematic (steps count), and physiological (ECG signal, beats per minute, blood pressure, and body temperature) parameters by using a wearable electrocardiogram (ECG) sensor worn by surgeons during the performance of minimally invasive surgical (MIS) procedures (conventional laparoscopy and robotic-assisted surgery -RAS-). For this purpose, data related to the surgeon's ECG sensor parameters were collected during forty-four MIS sessions conducted by eighteen surgeons with different levels of experience. Once the dataset was generated, two preprocessing techniques were applied: scaling and normalization. These two datasets were subsequently divided into two subsets: one containing 80% of the data for training and cross-validation, and the other containing 20% of data for testing. Several Artificial Intelligence (AI) techniques were applied to the training dataset to develop the predictive models. Finally, these models were validated on cross-validation and test datasets. PCA results showed that the physiological parameters (ECG signal, beats per minute, blood pressure and body temperature) were the most representative parameters of the ECG sensor. The predictive analysis results showed that XGBoost achieved the best results for the training dataset, while Multiple Linear Regression demonstrated the best performance for the cross-validation and test datasets combined with the scaled preprocessing technique, achieving the highest R2 coefficient and the lowest error for each parameter analyzed. The linear models were successfully validated on both cross-validation and test datasets, underscoring the potential for prediction of factors contributing to surgeon's health improvement during MIS procedures.

SB2-03

Can In-Context Learning enable Large Vision Language Models to detect ECG abnormalities?

Samuel Camba[1], Abraham Otero[2], Daniel García[2], Luciano Sánchez[1], and Nahuel Costa[1]
[1]Computer Science Department, University of Oviedo, Gijón, 33202, Asturias, Spain, [2]Department of Information Technologies, Institute of Technology, Universidad San Pablo-CEU, CEU Universities, Campus Montepríncipe, 28660, Spain, [3]Arrhythmia Unit, Hospital Universitario Central de Asturias, Oviedo, 33011, Asturias, Spain

Electrocardiogram (ECG) interpretation is essential for cardiac diagnosis, yet machine learning models often fail to adapt to realworld data, particularly when certain conditions, such as rare diseases, are underrepresented in training sets. In-Context Learning (ICL) has shown promise for adapting Large Language Models (LLMs) to new tasks using only a few labeled examples at inference time, but its potential in the domain of medical imaging remains largely unexplored. In this work, we investigate whether ICL can enable large Vision-Language Models (VLMs) to interpret ECG images. Our approach leverages VLMs to perform classification in data-constrained scenarios without parameter updates, offering a practical alternative when fine-tuning is not feasible. To enhance both interpretability and predictive performance, we incorporate a concept-based prompting framework inspired by Concept Bottleneck Models (CBMs). Specifically, the VLM is prompted to first predict clinically meaningful intermediate concepts, which then guide the final diagnosis, providing structured, interpretable reasoning. We evaluate our method on detecting Left Bundle Branch Block and Brugada syndrome. Results indicate that combining ICL with CBM-style prompting yields interpretable ECG analyses and promising diagnostic performance in data-limited settings. This suggests that VLMs could be a practical tool for extending machine learning to underrepresented clinical conditions without the need for extensive labeled data.

SB2-04

Novel automated tool for functional substrate assessment of the left atrium in patients with persistent AF using machine learning

Saman Golmaryami[1,2,3], Etel Silva Garcia[1,2,3], and Juan Fernandez Armenta[1,2,3]
[1]Instituto de Investigación e Innovación Biomédica de Cádiz (INiBICA), Spain, [2]Hospital Universitario Puerta del Mar (HUPM), Spain, [3]Universidad de Cádiz, Spain

Persistent atrial fibrillation (PsAF) presents high recurrence rates despite pulmonary vein isolation (PVI), highlighting the need to target the atrial substrate. Substrate characterization beyond the pulmonary veins using functional mapping can potentially unmask latent slow conduction areas. To improve substrate characterization in PsAF, an automated tool is developed to quantify key electrogram (EGM) features—duration, local activation time (LAT), number of deflections, and pacing-induced duration increase (Delta). The clinical data consist of over 1000 intracardiac EGMs of 5 PsAF patients that were recorded during baseline sinus rhythm and after three short-coupled extra stimuli. The features calculated by the tool achieved intra-class correlation coefficients of 0.98 for R3-LAT and 0.77 for R3-Duration, respectively, when compared to blinded expert clinicians' annotation, indicating reliable and accurate automated measurements. Furthermore, the extracted features from the paced EGMs showed strong discriminatory power in identifying hidden slow conduction substrates, achieving 93.6% balanced accuracy with a Support Vector binary classification model using a linear kernel.

SB2-05

A quantum machine learning approach to cardiopulmonary sound classification: QML for Cardiopulmonary Sound Classification

Sandra Ranilla-Cortina[1], Antonio J. Muñoz-Montoro[2], Elías F. Combarro[1], Sebastián García-Galán[2], and José Ranilla[1]
[1]Universidad de Oviedo, Spain, [2]Universidad de Jaén, Spain

Cardiopulmonary sound analysis is essential for the early detection and diagnosis of respiratory and cardiovascular diseases. This work explores the application of Quantum Machine Learning (QML) models, specifically Quantum Support Vector Machines (QSVM) and Quantum Neural Networks (QNN), for the classification of both lung and heart sounds. Leveraging MFCC-based features and dimensionality reduction techniques, we evaluate the performance of these models on two publicly available benchmark datasets. The experimental results indicate that QML models match or surpass their classical counterparts, particularly under constraints of limited training data and reduced feature sets. These findings underscore the potential of QML as a promising tool for efficient, accurate and unified analysis of cardiopulmonary acoustic signals in next-generation diagnostic systems.

SA3-01

A comparative analysis of message-level and user-level natural language processing approaches for early depression detection on social media

Carmen Rodríguez Jiménez[1], Miguel Rujas[2], Beatriz Merino-Barbancho[2], Maria Teresa Arredondo[2], Maria Fernanda Cabrera-Umpierrez[2], Kinda Khalaf[3], and Giuseppe Fico[2]
[1]Universidad Politécnica de Madrid, Avda Complutense 30, 28040 Madrid, Spain, [2]Life Supporting Technologies Research Group, Universidad Politécnica de Madrid, Avda Complutense 30, 28040 Madrid, Spain., [3]College of Medicine and Health Sciences, Khalifa University of Science and Technology, P.O. Box 127788 Abu Dhabi, UAE

Depression remains a key global mental health concern, with social media emerging as a promising source for early detection through user-generated content. This study presents a comprehensive comparison between message-level and user-level natural language processing (NLP) approaches for identifying early signs of depression on social media. Using the eRisk 2025 dataset, encompassing over 1.4 million Reddit posts from 3,061 users, we evaluated multiple machine learning and deep learning models, including Logistic Regression, Random Forest, XGBoost, and BERT. Results showed that while message-level models achieved high macro-level accuracy and AUC scores (up to 0.90), they struggled to reliably detect depressive messages, as evidenced by low F1-scores and precision in the depressive class. In contrast, user-level models, which aggregate information across multiple posts per user, demonstrated superior performance in identifying depressive users, with higher recall and F1-scores (up to 0.88 recall). These findings highlight the importance of analytical granularity in mental health detection tasks: user-level approaches offer a more robust and context-aware strategy for early identification of individuals at risk. The study demonstrates the potential of integrating artificial intelligence and NLP for proactive mental health monitoring, while also acknowledging challenges related to class imbalance, generalizability, and clinical validation. Future research should explore hybrid models, multi-platform data integration, real-time systems, and ethical frameworks to enhance practical applicability and societal impact.

SA3-02

A Weak Supervision Approach for Monitoring Recreational Drug Use Effects in Social Media

Lucía Prieto-Santamaría[1,2], Alba Cortés Iglesias[1], Claudio Vidal Giné[3], Fermín Fernández Calderón[4,5], Óscar M. Lozano[4,5], and Alejandro Rodríguez-González[1,2]
[1]Escuela Técnica Superior de Ingenieros Informáticos, Universidad Politécnica de Madrid, Spain, [2]Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Spain, [3]Asociación Bienestar y Desarrollo, Energy Control, Spain, [4]Clinical and Experimental Psychology Department, University of Huelva, Spain, [5]Research Center for Natural Resources, Health and the Environment, University of Huelva, Spain

Understanding the real-world effects of recreational drug use remains a critical challenge in public health and biomedical research, especially as traditional surveillance systems often underrepresent user experiences. In this study, we leverage social media (specifically Twitter) as a rich and unfiltered source of user-reported effects associated with three emerging psychoactive substances: ecstasy, GHB, and 2C-B. By combining a curated list of slang terms with biomedical concept extraction via MetaMap, we identified and weakly annotated over 92,000 tweets mentioning these substances. Each tweet was labeled with a polarity reflecting whether it reported a positive or negative effect, following an expert-guided heuristic process. We then performed descriptive and comparative analyses of the reported phenotypic outcomes across substances and trained multiple machine learning classifiers to predict polarity from tweet content, accounting for strong class imbalance using techniques such as cost-sensitive learning and synthetic oversampling. The top performance on the test set was obtained from eXtreme Gradient Boosting with cost-sensitive learning (F1 = 0.885, AUPRC = 0.934). Our findings reveal that Twitter enables the detection of substance-specific phenotypic effects, and that polarity classification models can support real-time pharmacovigilance and drug effect characterization with high accuracy.

SA3-03

Automated Classification of Ischemic Stroke Subtypes from Electronic Health Records Using Large Language Models

Alejandro Vásquez[1], Cristina Rubio-Escudero[2], and Germán Antonio Escobar-Rodríguez[1]

[1]Group for Research and Innovation in Biomedical Informatics, Biomedical Engineering, and Health Economy. Institute of Biomedicine of Seville, IBiS/"Virgen del Rocío" University Hospital /CSIC/University of Seville, Seville, Spain., [2]Department of Computer Languages and Systems, University of Sevilla, Spain.

This paper presents a study on the application of Large Language Models (LLMs) for the automatic classification of ischemic stroke subtypes from Electronic Health Records (EHRs). The study involves the creation and analysis of four datasets, an evaluation of model performance using standard classification metrics, and a comparison with expert manual labeling. Results suggest LLMs can offer promising outcomes even under computational constraints, paving the way for clinical automation and diagnostic support.

SA3-04

Concept Normalization in Psychiatry: Comparing Embedding and Lexical Methods for Spanish Clinical Text

Sergio Rubio-Martín[1], Arturo Crespo-Álvaro[1], María Teresa García-Ordás[1], Antonio Serrano-García[2], Clara Margarita Franch-Pato[2], and José Alberto Benítez-Andrades[1]
[1]ALBA Research Group, Department of Electric, Systems and Automatics Engineering, Universidad de León, Campus of Vegazana s/n, León, 24071, León, Spain, [2]Servicio de Psiquiatría, Complejo Asistencial Universitario de León (CAULE), 24008, León, Spain.

The automatic normalization of clinical entities is a critical step for enabling structured analysis of free-text medical records. This study proposes and evaluates four distinct retrieval strategies for normalizing entities extracted via Named Entity Recognition (NER) from Spanish psychiatric emergency notes, using Unified Medical Language System (UMLS) as the target terminology. A total of 768 annotated entities were mapped using MiniLM, Multilingual BERT, lexical matching, and the UMLS API. Results show that the API-based approach yields the best performance (Hit@3 = 56.8% and Hit@5 = 57.9%), effectively balancing accuracy and computational efficiency. Although embedding based methods such as MiniLM and Multilingual BERT often outperform traditional techniques in other domains, they showed only marginal improvements over simple lexical matching in this context, while incurring significantly higher computational costs. These findings suggest that in the psychiatric domain, where language is often ambiguous, embedding based approaches may fall short. Hybrid systems that combine fast retrieval with semantic reasoning are needed to capture the richness of psychiatric narratives.

SA3-05

Integrating Language Models and Network Embeddings to Uncover Hidden Relationships in Neuromuscular Diseases

Federico García-Criado[1], Jesús Pérez-García[1], Elena Rojano[1,2,3], Juan AG Ranea[1,2,3,4], and Pedro Seoane-Zonjic[1,2,3]

[1]Department of Molecular Biology and Biochemistry, University of Malaga, 29010 Malaga, Spain., [2]Institute of Biomedical Research in Malaga and platform of nanomedicine (IBIMA Plataforma BIONAND), 29071 Malaga, Spain., [3]Center for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III (ISCIII), 28029 Madrid, Spain., [4]Spanish National Bioinformatics Institute (INB/ELIXIR-ES), Instituto de Salud Carlos III (ISCIII), 28020 Madrid, Spain.

Neuromuscular diseases (NMDs) are a heterogeneous group of rare disorders that significantly impair motor function and quality of life. Their clinical and genetic variability makes systematic study and knowledge integration particularly challenging. We present a scalable, automated framework that combines natural language processing and network analysis to uncover hidden relationships among NMDs. Using Sentence Transformer models, we identified NMDs and their phenotypes in open-access documents. Then, we use this data to build NMD networks to infer new relations using embedding techniques and clustering approaches structuring a document corpus for 328 NMDs basis on disease type and associated to phenotypes and genes. These findings show the value of combining language models and network embeddings for large-scale rare disease analysis.

SB3-01

3D MTransINR: a 3D Modality Translation model based on Implicit Neural Representations

Maria Ysern[1] and Veronica Vilaplana[1]
[1]*Universitat Politècnica de Catalunya, Barcelona, Spain*

We present 3D MTransINR, a novel model for volumetric MRI modality translation based on Implicit Neural Representations (INRs). In clinical workflows, acquiring a complete set of MRI modalities is often impractical, motivating methods that can infer multiple complementary missing sequences from available ones. Our model tackles this predictive synthesis task by translating one or more source modalities into one or more targets within a unified, resolution-independent framework. 3D MTransINR combines a shared multilayer perceptron (MLP) with voxelwise bias modulation generated by a 3D U-Net, enabling anatomically consistent synthesis. We extend prior INR-based methods to full volumetric data and adversarial training, and evaluate our approach on the ProstateX and BraTS datasets. On ProstateX, 3D MTransINR outperforms a 3D Pix2Pix baseline in PSNR and SSIM across all targets and resolutions, showing strong robustness to resolution changes. On BraTS, it preserves finer structural details but is more sensitive to modality contrast, highlighting the challenges of generalisation across anatomies and intensity profiles.

SB3-02

Encoding the Spatio-Temporal Features of Rey-Osterrieth Complex Figure Strokes for Use in Deep Graph Networks

Pedro José Mulas Cámara[1], José Manuel Cuadra Troncoso[2], Mariano Rincón Zamorano[3], and Pedro Juan Tarraga López[4]
[1]Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain, [2]Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain, [3]Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain, [4]Universidad Castilla-La Mancha (UCLM), Albacete, Spain.

This work introduces a novel methodological approach for the analysis of the Rey-Osterrieth Complex Figure (ROCF), a task widely employed in neuropsychological research, by modeling the drawing process as a time series. The aim is to obtain a compact spatio-temporal representation that addresses the limitations of traditional static-image approaches. The process involves two stages: segmentation of the drawing into strokes to capture spatial and temporal information, and their encoding into feature vectors using deep learning models such as BiLSTM-Autoencoder and ST2Vec. This encoding enables dimensionality reduction and standardization of the spatio-temporal representation of strokes. Validation is currently based on reconstruction accuracy, showing that both architectures preserve essential spatial and temporal aspects: BiLSTM-Autoencoder achieves lower RMSE, while ST2Vec better retains spatial structure. These results should be regarded as a foundational step towards the integration of encoded strokes into graph structures (GNN), enabling future analyses of handwritten drawings from a spatio-temporal perspective. While the long-term vision of this project includes applications in cognitive assessment, the present contribution is methodological in nature and establishes the basis for future extensions to clinically validated populations.

SB3-03

Hybrid Morphology-Based Tumor Detection from Breast MRI Segmentation Masks

Sergio Botana[1], Paula Puerta González[2], Pablo García Marcos[1,3], Sara Fernández Arias[1], Rebeca Oliveira Suárez[2], Guillermo Lorenzo[7,8], Héctor Gómez[4,5,6], Covadonga del Camino[9], Víctor M. González[2,3], and Angel Rio-Alvarez[1,3]

[1]Computer Science Department, University of Oviedo, Spain, [2]Electrical Engineering Department, University of Oviedo, Spain, [3]Biomedical Engineering Center, University of Oviedo, Spain, [4]School of Mechanical Engineering, Purdue University, 585 Purdue Mall, West Lafayette, IN 47907, USA, [5]Purdue Center for Cancer Research, Purdue University, West Lafayette, IN, USA, [6]Weldon School of Biomedical Engineering, Purdue University, 206 S. Martin Jischke Drive, West Lafayette, IN 47907, USA, [7]Group of Numerical Methods in Engineering Department of Mathematics Campus de Elviña s/n, 15071 A Coruña, Galicia, Spain, [8]Oden Institute for Computational Engineering and Sciences, The University of Texas at Austin, 201 E. 24th Street, Austin, TX 78712-1229, USA, [9]Radiodiagnostic Service Oviedo Central University Hospital (HUCA), Oviedo, Spain.

Magnetic resonance imaging (MRI) is widely used to diagnose breast cancer. However, the high sensitivity of dynamic contrast-enhanced MRI (DCE-MRI) comes at the cost of increased complexity in image interpretation and segmentation, which requires advanced techniques to help radiologists delineate tumor boundaries accurately and efficiently. This work proposes an automated methodology that reduces the reliance on manual annotations by integrating unsupervised clustering with supervised classification to identify the tumor cluster within masks derived from clustering techniques using only morphological characteristics, thus facilitating rapid, accurate, and reproducible tumor detection in breast magnetic resonance imaging. The approach consists of an initial segmentation using Fuzzy C-Means followed by classification using supervised machine learning models. A comprehensive feature selection analysis that identifies the most informative descriptors for tumor classification, resulting in a classification accuracy of 99.44%, with consistent improvements across all evaluated metrics. The proposed approach mitigates annotation dependency while ensuring interpretability, representing a practical tool for clinical deployment.

SB3-04

Multimodal Posterior Sampling-based Uncertainty in PD-L1 Segmentation from H&E Images

Roman Kinakh[1], Gonzalo R. Rios-Muñoz[1,2], and Arrate Muñoz-Barrutia[1]
[1]Universidad Carlos III de Madrid, Leganés, Madrid, Spain, [2]Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

Accurate assessment of PD-L1 expression is critical for guiding immunotherapy, yet current immunohistochemistry (IHC) based methods are resource-intensive. We present nnUNet-B: a Bayesian segmentation framework that infers PD-L1 expression directly from H&E-stained histology images using Multimodal Posterior Sampling (MPS). Built upon nnUNet-v2, our method samples various model checkpoints during cyclic training to approximate the posterior, enabling accurate segmentation and epistemic uncertainty estimation via entropy and standard deviation. Evaluated on a dataset of lung squamous cell carcinoma, our approach achieves competitive performance against established baselines with a mean Dice Score and a mean IoU of 0.805 and 0.709, respectively, while providing pixel-wise uncertainty maps. Uncertainty estimates show a strong correlation with segmentation error, while calibration can be further optimized. These results suggest that uncertainty-aware H&E-based PD-L1 prediction is a promising step toward scalable and interpretable biomarker assessment in clinical workflows.

SB3-05

Integrating Radiomics and Deep Learning for Breast Lesion Classification

Eduardo Almeda Luna[1,2], José María Luna[1,2], Rosa Sicila[3], Valerio Guarrasi[3], Paolo Soda[3], and Sebastián Ventura[1,2]
[1]Department of Computer Science and Numerical Analysis, Andalusian Research Institute in Data Science and Computational Intelligence (DaSCI), Spain, [2]Maimonides Biomedical Research Institute of Cordoba, IMIBIC, University of Cordoba, Córdoba, 14071, Spain, [3]Unit of Artificial Intelligence and Computer Systems, Department of Engineering, Università Campus Bio-Medico di Roma, Rome, 00128, Italy

Breast cancer is a leading cause of death in women, and early, accurate diagnosis is vital. While mammography is standard, distinguishing benign from malignant lesions is still difficult. Radiomics and deep learning show promise in enhancing diagnostic accuracy. This study combines radiomic features and convolutional neural networks (CNNs) to improve the classification of lesions in mammography. A comparative evaluation is performed across neural networks trained only with radiomic features (around 1,000 features per lesion, extracted using PyRadiomics), standard CNNs, and hybrid models that combine CNNs and radiomic features. A subset of the Chinese Mammography Database consisting of 332 breast lesions (180 benign, 152 malignant), including craniocaudal and mediolateral oblique views, were used. Experimental results demonstrated that CNN-based models (F1-Score: 0.806±0.051) outperformed networks using only radiomic features (0.723±0.037). Additionally, hybrid methods produced the best results (0.822±0.026). These findings underscore the potential of combining CNNs and radiomics in breast cancer diagnosis, reducing the need for biopsies.

SB4-01

Machine Learning for MRI-Based Classification of Treatment Response in Diffuse Gliomas

María Yanshuang Martín Moreira[1], Ana María Pérez Martín[1], José-Javier Serrano-Olmedo[2,5,6], Ángel Luis Álvarez [4] and Oscar Casanova-Carvajal [2,3]
[1]Universidad Alfonso X El Sabio, Madrid 28691, Spain., [2]Centro de Tecnología Biomédica, Campus de Montegancedo, Universidad Politécnica de Madrid, 28040 Madrid, Spain, [3]Departamento de Ingeniería Eléctrica, Electrónica automática y Física aplicada, Escuela Técnica superior de ingeniería y diseño industrial ETSIDI, Universidad Politécnica de Madrid, 28040 Madrid, Spain, [4]Escuela de Ingeniería de Fuenlabrada, Universidad Rey Juan Carlos, 28922 Madrid, Spain, [5]Centro de Investigación Biomédica en Red para Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Madrid, 28029, Spain, [6]Departamento de Tecnología Fotónica y Bioingeniería, ETSI Telecomunicaciones, Universidad Politécnica de Madrid, Madrid, 28040, Spain

Diffuse gliomas are currently challenging to classify in the field of brain tumours due to their clinical and imaging characteristics. This study aims to develop and train a patient classification system based on the study of various parameters to support personalized treatment planning. The use of MATLAB allows us to analyse, segment, and process all data to extract relevant features associated with glioma characteristics. Classification was performed by considering Radiomics (such as texture and shape), biomolecular makers, and clinical information from the patient, in order to predict therapeutic response. Experimental results show that the system is able to distinguish subgroups of patients with moderate but consistent accuracy, demonstrating the feasibility of applying computational image analysis in combination with clinical and molecular variables. The integration of radiomics, biomolecular, and clinical information underscores the potential of such an approach as a decision-support tool in neuro-oncology. Overall, the findings suggest that multimodal classification systems represent a promising pathway toward improving the accuracy of glioma diagnosis and advancing personalized treatment planning, ultimately contributing to better patient outcomes.

SB4-02

Segmentation of Arabidopsis Apical Stem Cells via a Dual Deep Learning Approach

Guillermo Rey-Paniagua[1], Dariusz Lachowski[1], and Arrate Muñoz-Barrutia[1,2]
[1]Universidad Carlos III de Madrid, Leganés, Madrid, Spain, [2]Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Spain

We introduce a deep learning pipeline for segmenting apical stem cells in three-dimensional confocal microscopy volumes of Arabidopsis thaliana. By integrating pre-trained 2D and 3D U-Net models from the BioImage Model Zoo, our method combines in-plane boundary sharpness with volumetric continuity. The workflow encompasses parallel preprocessing, dual-model inference, logical fusion, 3D reconstruction, and membrane-aware post-processing to extract key morphometric features, including cell counts and volumes. Evaluated over 22 timepoints, our pipeline achieves a mean Dice similarity coefficient (DSC) of 0.886 ± 0.002. Passing–Bablok regression on cell counts yields a slope of 0.885 ($p = 0.573$), surpassing the standalone 2D U-Net (0.539) and 3D UNet (0.782). Volume estimates exhibit a slope of 0.754 ($p = 0.107$), with smaller cell volumes but substantially fewer cell-merging artifacts compared to the baselines (0.738 and 0.930, respectively). These results highlight the advantages of model fusion for robust, biologically meaningful segmentation. The complete pipeline and source code are publicly available at: https://github.com/GolpedeRemo37/Arabidopsis_DualDL.

SA4-01

Molecular Machine Learning Using Euler Characteristic Transforms

Victor Toscano-Duran[1], Florian Rottach[2,3], and Bastian Rieck[4]
[1]Department of Applied Mathematics I, University of Seville, Seville, Spain, [2]Central Data Science, Boehringer Ingelheim GmbH, Biberach/Riss, Germany, [3]School of Medicine, University of Tübingen, Germany, [4]AIDOS Lab, University of Fribourg, Fribourg, Switzerland

The shape of a molecule determines its physicochemical and biological properties. However, it is often underrepresented in standard molecular representation learning approaches. Here, we propose using the Euler Characteristic Transform (ECT) as a geometrical-topological descriptor. Computed directly from molecular graphs constructed usinghandcrafted atomic features, the ECT enables the extraction of multiscale structural features, offering a novel way to encode molecular shape in the feature space. We assess the predictive performance of this representation across nine benchmark regression datasets, all centered around predicting the inhibition constant Ki. In addition, we compare our proposed ECT-based descriptor against traditional molecular representations and methods, such as molecular fingerprints/descriptors and graph neural networks (GNNs). Our results show that our ECT-based representation achieves competitive performance, ranking among the best-performing methods on several datasets. More importantly, combining our descriptor with established representations, particularly with the AVALON fingerprint, significantly enhances predictive performance, outperforming other methods on most datasets. These findings highlight the complementary value of multiscale topological information and its potential for being combined with established techniques. Our study suggests that hybrid approaches incorporating explicit shape information can lead to more informative and robust molecular representations, enhancing and opening new avenues in molecular machine learning. To support reproducibility and foster open biomedical research, we provide open access to all experiments and code used in this work.

SA4-02

Sharing Patterns Between Proteins and Exploring Possible Drug Interactions

Paloma Tejera-Nevado[1] and Alejandro Rodríguez-González[1,2]
[1]Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, Spain,
[2]ETS Ingenieros Informáticos, Universidad Politécnica de Madrid, Madrid, Spain

Due to the nature of amino acids that make up proteins, it is possible to detect small patterns within their sequence. Some of these sequences have been identified as motifs, characterized by defined signatures that are shared across different proteins. In this study, patterns are defined as short amino acid sequences that are commonly found among proteins involved in various types of cancer. The research focuses on the localization of such patterns in ALDH2, chosen as a reference protein because it is known to be inhibited by disulfiram. Disulfiram, originally used to treat alcohol use disorder (AUD), has been repurposed for the treatment of lung cancer. Analyzing shared patterns can help identify common protein structures and may assist in locating potential binding cavities and interaction sites, while also improving our understanding of how related or structurally similar proteins might be affected by certain drugs. This approach provides a new perspective on drug specificity and targeting, highlighting three achievements: identifying conserved amino acid motifs in cancer-related proteins, a case study on ALDH2 inhibition by disulfiram, and showing how motif analysis can guide drug repurposing.

SA5-01

A Framework for Evaluating the Stability of Learned Representations in Biologically-Constrained Models in Single-Cell: Evaluating the Stability of Biologically-Informed Models

Sara Fernandez-Malvido, Alberto Esteban-Medina, Pelin Gundogdu, Joaquin Dopazo, Isabel A. Nepomuceno-Chamorro, and Carlos Loucera
[1]*Universidad Carlos III de Madrid, ROR: https://ror.org/03ths8210. Computer Science and Engineering Department. Applied Artificial Intelligence Group, Spain*

This work addresses the instability and poor interpretability of computational models in single-cell RNA sequencing (scRNA-seq) analysis. We propose a generalizable framework to evaluate the stability of any model that generates pathway-level scores, applying it to both biologically-constrained variational autoencoders (iVAEs) and alternative graph-based methods like PathSingle. The central contribution is a modular and reproducible workflow, orchestrated with Pixi, Prefect and Ray, that automates the systematic comparison of different models across multiple random seeds. The stability of the learned representations (pathway activities) was assessed using metrics for clustering coherence (Adjusted Mutual Information, AMI) and consistency across runs (hyperbolically weighted Kendall's Tau, $w\tau$). Our framework revealed that iVAEs informed by biological priors are significantly more stable and produce more meaningful cell groupings than randomly connected counterparts, therefore indicating the importance of being informed by meaningful biological entities. While the PathSingle model demonstrated a marginally superior consistency, the informed VAEs offered a better balance between stability and clustering performance. This work provides a robust methodology for assessing diverse pathway scoring models, promoting the development of more reliable and interpretable tools for single-cell analysis.

SA5-02

An Interpretable Graph Neural Network for Multi-Omics Data Integration and Biomarker Discovery

Alberto Labarga[1]
[1]*Universidad Pública de Navarra, Spain*

The integration of heterogeneous multi-omics data remains a critical challenge in computational biology, essential for unraveling the complex molecular underpinnings of diseases. Existing methods often struggle to incorporate prior biological knowledge effectively or provide interpretable results. We present BioMGNN v2, an end-to-end, supervised deep learning framework designed to address these challenges by extending our previous BioMGNN architecture. BioMGNN v2 constructs a heterogeneous biomedical knowledge graph that embeds patient-specific multi-omics data within a network of signed, directed biological interactions. It then leverages a heterogeneous graph neural network with relation-specific attention and integrates a complementary patient–patient graph branch, enabling robust representation learning across both biological entities and clinical samples. Self-supervised pretraining with contrastive and masked modeling objectives further enhances generalization, while gated cross-omics fusion and multi-task learning improve classification accuracy and biological relevance. To ensure interpretability, BioMGNN v2 incorporates graph explainers with stability selection to identify reproducible and statistically rigorous biomarker modules, complemented by uncertainty estimation for clinical reliability. We demonstrate the superior performance of our algorithm in patient classification and its ability to uncover biologically plausible and interpretable biomarkers, positioning it as a powerful and trustworthy tool for advancing precision medicine.

SA5-03

Evaluation of Deep Clustering methods on high dimensional tabular biomedical data

Ruben E. Munoz-Cabrera[1], Manuel Campos[1,2], and Jose M. Juarez[1,2]
[1]MedAI Lab, University of Murcia, Spain, [2]Murcian Institute of Biomedicine, Spain

Clustering is a fundamental task in biomedical research, particularly for enabling stratification of patients. Traditional clustering techniques often struggle in this domain due to the vast number of attributes to be considered and the presence of complex, non-linear data structures. Deep Clustering (DC) methods, which combine neural networks with classical clustering techniques, have emerged as a promising alternative to address this challenge. To this day, most DC research has focused on images or text, however for many biomedical problems data is in a tabular format. This work proposes and evaluates a systematic framework for assessing clustering methods on high-dimensional tabular data. The framework consists of four stages: defining the experimental setup, selecting and preprocessing datasets, implementing and executing experiments, and performing evaluation and comparative analysis. Experiments were conducted on four genomic datasets, comparing traditional clustering with and without dimensionality reduction to DC approaches (e.g. DEC, VaDE or TableDC). Results indicate classical techniques perform well in structured scenarios while DC methods show advantages in complex or imbalanced datasets. This work highlights the importance of selecting clustering techniques based on dataset characteristics. It demonstrates the potential of DC for high dimensional data.

SA5-04

Integrative analysis of breast cancer using multi-omics latent representations

Yasmine Lakouifat Darkaoui[1], Beatriz Pontes[1], and Belén Vega Márquez[1]
[1] *Computer Languages and Systems Department, University of Seville*

This study presents a comparative analysis of three dimensionality reduction techniques—MOFA (Multi-Omics Factor Analysis), IntNMF (Integrative Non-negative Matrix Factorization), and VAE (Variational Autoencoder)—applied to breast cancer multi-omics data from the TCGA-BRCA project. The integration of omics layers—including gene expression, protein expression, copy number variations, and mutations—was combined with key clinical variables to evaluate the performance of latent representations in both classification and clustering tasks. Major challenges such as high dimensionality and severe class imbalance were addressed through oversampling and undersampling strategies. Each method was evaluated for its effectiveness in predicting clinical outcomes and identifying meaningful molecular patterns. MOFA offered biologically interpretable and stable representations, IntNMF produced compact structures, and VAE yielded well-separated latent spaces. Enrichment analysis confirmed the relevance of extracted features, reinforcing the utility of latent factor models for robust multi-omics integration in breast cancer research.

SA5-05

Optimization of Denoising Autoencoders with Progressive Learning Strategies for scRNA-seq Data

Francisco Javier Franco-Ruiz[1], Sara Fernandez-Malvido[1], Belen Vega-Marquez[1], and Isabel A. Nepomuceno-Chamorro[1]
[1]*Dpto. de Lenguajes y Sistemas Informáticos, University of Seville, Sevilla, Spain*

Advances in single-cell RNA sequencing (scRNA-seq) technologies have revolutionized cancer research by enabling detailed characterization of cellular heterogeneity [6]. However, predicting drug sensitivity at single-cell resolution remains challenging due to the scarcity of annotated data, high dimensionality, and technical noise. In this work, we reproduce and enhance the scDEAL framework, a deep transfer learning model that integrates bulk and single-cell transcriptomic data for drug response prediction. The enhancements focus on optimizing Denoising Autoencoders (DAE) and incorporating progressive training strategies. Specifically, we introduce Gene Prioritization Regularization (GPR) to emphasize biologically relevant genes, implement Curriculum Learning to gradually increase task complexity during training, and apply direct filtering of highly variable genes to reduce dimensionality. Experiments conducted on publicly available datasets from GDSC, CCLE, and GEO demonstrate that filtering the top 20% most variable genes leads to significant improvements in predictive performance, achieving an F1-score of 0.9641 and an AUC of 0.9549, while reducing computational costs by 80%. These results highlight the importance of feature selection and progressive training strategies for enhancing drug response prediction from scRNA-seq data.

SA6-01

Integrative Analysis of Gene Co-Expression Networks and Biclustering for Cancer Biomarker Discovery

Marc Ríos Cadenas[1], Aurelio López-Fernández[1], Iván Segura-Carmona[1], Juan A. Ortega[2], and Francisco A. Gómez-Vela[1]
[1] Intelligent Data Analysis Group (DATAi), Universidad Pablo de Olavide, Ctra. Utrera km. 1, ES-41013 Seville, Spain, [2]Computer Science Department, Universidad de Sevilla, ES-41012 Seville, Spain.

Soft tissue sarcomas (STS) are rare and heterogeneous cancers with limited biomarkers. Comprehending their underlying mechanisms is crucial for identifying biomarkers that improve diagnosis and facilitate targeted therapy. This paper applies two independent strategies—gene co-expression network (GCN) analysis and biclustering—to identify biomarkers in uterine leiomyosarcoma (ULMS), using one RNASeq and one microarray dataset. GCN analysis uncovered globally coexpressed hub genes like FOXM1, E2F1, MYBL2, and PITX1, linked to mitosis and DNA replication. In contrast, biclustering identified local coexpression patterns, including FOXM1 and HLF as transcriptional regulators, and other genes such as CCNB1, POLQ, and TRIP13 involved in genomic stability. Notably, FOXM1 was detected by both methods, reinforcing its relevance. While GCN highlights global regulatory roles, biclustering captures condition-specific signals. Their complementarity enhances biomarker discovery and contributes to understanding ULMS transcriptional architecture.

SA6-02

Metabopathia: Enhancing disease mechanism understanding through mechanistic integration of transcriptomic and metabolic data

Kinza Rian[1,4], Isabel A. Nepomuceno-Chamorro[4], Joaquin Dopazo[1,2,3], and Daniel López-López[1,2,3]

[1] Andalusian Platform for Computational Medicine, Andalusian Public Foundation Progress and Health-FPS, Seville, Spain, [2]Institute of Biomedicine of Seville (IBiS), University Hospital Virgen del Rocío / CSIC / University of Seville, Seville, Spain., [3]Centro de Investigación Biomédica en Red en Enfermedades Raras (CIBERER), ISCIII, Madrid, Spain., [4]Department of Lenguajes y Sistemas Informáticos, University of Seville, Seville, Spain

Multifactorial diseases arise from complex interactions among various biological systems. These interactions represent intricate molecular mechanisms leading to pathological conditions. Studying these mechanisms requires dedicated analytical methods capable of integrating diverse omics data to model disease progression. Traditional analytical methods focus on the integration of individual genetic, transcriptomic, or metabolic datasets which are insufficient for capturing this complexity. Integration of one layer of omics data may lead to significant gaps in our understanding of complex diseases. To address this limitation, we present Metabopathia, an extension of HiPathia, as a computational tool that integrates multi-omics data—including transcriptomics and metabolomics—into mechanistic models of signaling and metabolic pathways. The method leverages curated biological knowledge from the Kyoto Encyclopedia of Genes and Genomes (KEGG), which provides the pathway structures used to model these processes. By doing so, Metabopathia enables accurate measurement of changes in the activity of cellular signaling cascades. This not only supports highthroughput estimation of functional cellular profiles, but also allows the simulation of how genetic mutations that reduce or eliminate gene function, along with metabolic disruptions, affect cellular processes. The application of Metabopathia to complex diseases such as cancer has demonstrated its efficacy in identifying critical sub-pathway alterations that drive disease progression. By incorporating metabolomic data, Metabopathia improves the mechanistic understanding of multifactorial diseases and informs potential therapeutic strategies tailored to disease-specific molecular contexts.

SA6-03

Predicting obesity-related phenotypes from the human gut microbiome using machine learning

Ivon-Teresa Sánchez-Cárdenas[1] and Martha-Ivon Cárdenas[2]
[1]Universitat Pompeu Fabra (UPF), Barcelona, Spain., [2]Soft Computing Research Group (SOCO) at Intelligent Data Science and Artificial Intelligence (IDEAI-UPC) Research Centre, Universitat Politècnica de Catalunya (UPC Barcelona Tech), Jordi Girona 1-3, 08034 Barcelona, Spain

The gut microbiome is increasingly recognized as a contributing factor in host metabolism, immune function, and energy regulation. Although previous studies have used machine learning (ML) to predict obesity from microbiome data, the present approach emphasizes model interpretability and reproducibility using SHAP-based feature ranking in publicly available datasets. In this study, interpretable machine learning techniques were applied to publicly available 16S rRNA amplicon data (ENA accession ERP003612) corresponding to individuals from the MetaHIT cohort, with the aim of predicting obesity-associated phenotypes from operational taxonomic unit (OTU) profiles. Several classifiers were evaluated (RandomForest, CatBoost, LightGBM, XGBoost, GBM, Logistic Regression, Naive Bayes), class imbalance was addressed via SMOTE within a nested 5×5 cross-validation framework, and model decisions were interpreted using SHAP. RandomForest achieved the highest ROC AUC (0.678 ± 0.039), with CatBoost close behind (0.674±0.056). SHAP analysis identified top OTUs mostly from the phyla Firmicutes and Bacteroidetes, including genera such as Clostridium, Coprobacillus, Odoribacter, and Blautia, and families like Clostridiaceae, Lachnospiraceae, and Ruminococcaceae. This study demonstrates that the application of machine learning to publicly available datasets can reveal predictive microbial signatures that might be overlooked by conventional analytical approaches. The proposed framework facilitates systematic hypothesis generation, selects relevant taxa for subsequent experimental validation, and provides reproducible insights into microbiome patterns associated with obesity.

SA6-04

Triclustering-Based Analysis of Circadian Gene Expression Patterns

Javier Hiruelo-Pérez[1], José M. García-Heredia[2], David Gutiérrez-Avilés[1], and María Martínez-Ballesteros[1]

[1] Department of Computer Languages and Systems, University of Seville, ES-41012 Seville, Spain., [2]Department of Vegetal Biochemistry and Molecular Biology, University of Seville, ES-41012 Seville, Spain

The circadian cycle orchestrates physiological and molecular processes through rhythmic gene expression. Uncovering relational patterns that link specific genes, experimental conditions, and time points is essential to understand the regulatory architecture of these rhythms. This study applies TriGen, a triclustering algorithm that identifies coherent gene–condition–time modules in circadian transcriptomic datasets. The analysis involved data preprocessing, tricluster extraction with TriGen, and functional enrichment analysis. TriGen uncovers biologically interpretable modules enriched in core clock components, as well as immune-related pathways, illustrating the potential of triclustering to reveal complex, condition-dependent circadian relationships. Across ten independent experiments, the method recovered up to seventy two percent of a curated set of literature-derived circadian genes, capturing both canonical clock regulators and auxiliary modulators involved in metabolism, immune signaling, and stress response. Functional enrichment analysis revealed significant associations with mitochondrial function, transcriptional regulation, neuronal signaling, and immune defense, reflecting the multifaceted systemic impact of circadian regulation. These results highlight the value of triclustering for dissecting high-dimensional temporal transcriptomic data and provide a framework to uncover novel regulatory modules that can guide targeted experimental validation.

SA6-05

Disease Understanding and Visualization of Drug Repurposing Hypotheses

Lucía Prieto-Santamaría[1,2], Belén Otero-Carrasco[2], Andrea Álvarez-Pérez[2], David Juste Urraca[2], and Alejandro Rodríguez-González[1,2]
[1]Escuela Técnica Superior de Ingenieros Informáticos, Universidad Politécnica de Madrid, Spain., [2]Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Spain

The increasing availability of biomedical data has opened new opportunities to rethink how diseases are modeled and how drug repurposing hypotheses can be systematically generated. In this context, we present a web-based platform that (i) provides access to previously integrated heterogeneous biomedical data and (ii) enables interactive visualization of disease understanding relationships, as well as of drug repurposing hypotheses. The platform constructs multilayer disease networks, visualizes disease relationships, and implements six complementary computational repurposing methods, including more naive methods, network-based analysis, and artificial intelligence-based approaches. It also provides the web services needed to access the whole set of data. This paper describes the underlying data sources and methodologies and demonstrates its capabilities through a representative visualization case.

SB5-01

Characterising Continuous Glucose Monitoring Using Topological Data Analysis

Ana Lopez-Caballero[1], Miguel A. Meroño[2], and Jose M. Juarez[1,3]
[1]MedAI Lab, University of Murcia, Spain., [2]Faculty of Mathematics, University of Murcia, Spain., [3]Murcian Institute of Biomedicine, Spain

Topological Data Analysis (TDA) has increasingly gained recognition as a powerful and versatile mathematical framework aimed at extracting high-level structural features from complex datasets. Among the various tools that TDA offers, persistent homology has emerged as one of the most fundamental and widely adopted. However, clinical information is often presented in the form of time series. The present work focuses on applying TDA to data derived from continuous glucose monitoring (CGM) devices worn by diabetic patients. The study explores the potential of using persistent homology to transform CGM time series into meaningful topological summaries that may aid in distinguishing between patients diagnosed with Type 1 diabetes and those with Type 2. The experimental findings indicate that the exclusive reliance on topological information can discriminate between type 1 and type 2 diabetes patients. Nonetheless, the study is preliminary and limited to basic time series classification techniques.

SB5-02

Exploratory Computational Phenotyping of Hyposalivation Etiologies in Women

Kristina Lacasta[1], María J. Rus[1], Angela de la Cruz Gándara Alvarez[1], Cristiane Cantiga-Silva[1], Virginia Moreira Navarrete[2], Carmen Dominguez Quesada[2], Jose Javier Perez Venegas[2], Juan Antonio Ortega[3], and Aurea Simon-Soro[1,4]

[1] Department of Stomatology, Dental School, University of Seville, 41009 Seville, Spain., [2]Rheumatology Department, Virgen Macarena University Hospital (HUVM), 41009 Seville, Spain., [3]Computer Science Department, University of Seville, 41012 Seville, Spain., [4]Andalusian Center for Molecular Biology and Regenerative Medicine (CABIMER) Sevilla Spain.

Hyposalivation in women presents diagnostic challenges due to overlapping etiologies including medications, hormonal changes and autoimmune conditions. This exploratory study investigates whether computational approaches can identify distinct hyposalivation phenotypes in women. We analyzed 196 women across 40 clinical variables including demographics, medications, dental health, and salivary flow measurements using unsupervised machine learning. K-means clustering with PCA preprocessing tentatively identified three clusters: medication and lifestyle-associated hyposalivation (n=141, 72%), optimal salivary function (n=18, 9%), and severe hyposalivation (n=37, 19%). Salivary flow parameters and dental health variables were primary cluster drivers. Results provide preliminary evidence that hyposalivation may comprise distinct phenotypes, offering proof-of-concept for computational phenotyping approaches that warrant validation in independent cohorts.

SB5-03

Predictive Modeling with Machine Learning in Patients with Locally Advanced Rectal Adenocarcinoma Undergoing Neoadjuvant Treatment

Miguel Pagés García[1], Inmaculada Rincón Pérez[2], Juan Antonio Ortega Ramírez[1], José Luis López-Guerra[2], and Francisco Antonio Gomez Vela[3]
[1] Higher Technical School of Computer Engineering, University of Seville, 41012 Seville, Spain. [2]Department of Radiation Oncology, University Hospital Virgen del Rocío, 41013 Seville, Spain., [3]Intelligent Data Analysis Group (DATAi), Universidad Pablo de Olavide, Ctra. Utrera km. 1, ES-41013 Seville, Spain.

Colorectal cancer is one of the diseases with the highest rate of death worldwide, so it is necessary to improve its diagnosis and treatment. In this sense, new artificial intelligence techniques have emerged as a new way to understand the intrinsic mechanisms of the disease. In the present study, data analysis and machine learning techniques have been applied with the aim of identifying clinical and biological patterns that allow for optimizing patient stratification and improving the treatments administered. Through data cleaning and analysis, distinct patient profiles have been defined, clearly differentiating those with a poorer prognosis from those with a better outlook, thanks to clustering algorithms and the interpretation of the resulting clusters supported by the LIME technique. In addition, a supervised analysis has been performed, also supported by LIME to improve interpretability, which revealed the most influential variables in predicting key clinical events such as survival, tumor progression, and the pathological response of the tumor to treatment. These results highlight the importance of monitoring molecular markers and hematological parameters, as well as ensuring the proper management of treatment timelines, laying the foundation for improving clinical decision-making and advancing toward more personalized medicine.

SB5-04

Reduction of a Neuropsychological Test Battery Using Machine Learning Methods

Alba Gómez-Valadés Batanero[1], Rafael Martínez Tomás[1], Mariano Rincón[1], and Pedro Juan Tarraga López[2]

[1]Universidad Nacional de Educación a Distancia, 28040 Madrid, Spain., [2]Universidad de Castilla-La Mancha, Albacete, Castilla-La Mancha, Spain

To reach a broader population during Mild Cognitive Impairment population screenings, there is a need for neuropsychological test batteries that are quick to administer and evaluate. To achieve that, all test scores must be informa-tive and non-redundant. This study used machine learning methods to ana-lyze the test scores in a neuropsychological database in terms of relevance and collinearity. We examined a database of 520 assessments with diagnoses of either healthy or MCI, evaluated using a neuropsychological test battery consisting of fourteen tests, from which we will use twenty-three test scores. First, we assessed both the correlation between test scores and the diagnostic dependence of each score. This led us to define four groupings of test scores (very high relevance, very high and high relevance, all except low-relevance, and all test scores) that exclude redundant test scores within each group, and a fifth group that includes all test scores, even redundant ones, as a baseline. Those test score groups were analyzed using well-established machine learn-ing methods. Based on the machine learning analysis, the fifth group had the best performance, likely because of the redundant test scores. Among the groupings that exclude redundant test scores, we found that the most promis-ing group was the one with all test scores except the low-relevant ones, though it was almost tied in performance with the group containing all test scores.

SB5-05

Accurate Fall Risk Prediction in Older Adults: Integrating Sensor and Clinical Data Through Machine Learning

Ana González Castro[1], José Alberto Benítez-Andrades[2], and Raquel Leirós-Rodríguez[3]
[1] Nursing and Physical Therapy Department, Universidad de León, 24402 Ponferrada, Spain., [2]ALBA Research Group, Department of Electric, Systems and Automatics Engineering, Universidad de León, 2401 León, Spain., [3]SALBIS Research Group, Nursing and Physical Therapy Department, University of León, 24402 Ponferrada, Spain

Accurate identification of older adults at high risk of falling is essential to prevent injuries and implement effective interventions. This study evaluated the performance of several machine learning models in predicting fall risk using both accelerometric data from wearable sensors and non-accelerometric data including demographic, functional, and clinical variables. A dataset comprising 146 older women was used to train and compare seven algorithms: random forest, XGBoost, AdaBoost, LightGBM, Bayesian ridge regression, decision trees, and support vector regression. Predictive accuracy was assessed using mean squared error, mean absolute error, and the coefficient of determination. Models trained with combined accelerometric and non-accelerometric data consistently outperformed those using a single data source. XGBoost achieved the lowest MSE, while Bayesian ridge regression reached the highest R^2, highlighting its superior explanatory capability. In contrast, support vector regression exhibited poor predictive performance. Non-accelerometric variables, especially age and comorbidities, emerged as major predictors, whereas accelerometric data alone yielded limited accuracy. The integration of multiple data types significantly improved model robustness and clinical applicability. These findings underscore the value of multi-source data fusion in enhancing fall risk prediction among older adults and support the implementation of hybrid machine learning models in clinical settings.

SB6-01

Time Series Models: Application to RespiCast (ECDC Respiratory Diseases Forecasting Hub)

Lucía Maza[1], Andrés M. Alonso[2], and Carolina García-Martos[1]
[1]Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, C/José Gutiérrez Abascal, 2, 28006, Madrid (Spain)., [2]Departamento de Estadística and Instituto Flores de Lemus, Universidad Carlos III de Madrid, Getafe (Spain).

In this paper several univariate and multivariate (more specifically dimensionality reduction techniques, such as the Dynamic Factor Model, DFM) time series approaches and different training strategies for these models are proposed in order to compute forecasts for the number of cases per 100,000 inhabitants of Acute Respiratory Infections (AcRI) in several countries of the European Union (EU) whose historical datasets are available from RespiCast. All the proposals are used to compute weekly forecasts from 2020 up to spring 2025 and carefully compared using an adequate Analysis of Variance (ANOVA). Here, the forecasting horizon is extended up to 22 weeks, although in RespiCast the models provided by all the participants are 4 weeks maximum. As a summary, it can be stated that univariate ARIMA models accounting for seasonality are better for the short-term, and DFM produces more accurate forecasts in the long run. Additionally, the combination of forecasts here built with our best reduces the forecasting error of the baseline provided by RespiCast about 50% for the short-term. In some recent periods (first two months in 2025) our approach beats (in accuracy terms) most forecasts submitted by other teams (ranking 1, 2 or 3 out of 10 teams). For the rest of the span of time considered the average ranking is 4. It should be taken into account that just our DFM approach is capable of computing relatively accurate forecasts for longer forecasting horizons (up to 22 weeks) than those considered in RespiCast (just up to 4 weeks).

SB6-02

Synthetic Disease Trajectories for Predicting Chronic Kidney Disease Post-Diabetes Onset

Víctor M. de la Oliva Roque[1], David P. Kreil[2], Joaquín Dopazo[1,3], Francisco Ortuño[4], and Carlos Loucera[1,3]

[1] Andalusian Platform for Computational Medicine, Andalusian Public Foundation Progress and Health-FPS, Seville, Spain, [2]Institute of Molecular Biotechnology, Department of Biotechnology and Food Science, BOKU University, Vienna, Austria, [3]Institute of Biomedicine of Seville (IBiS), University Hospital Virgen del Rocío/CSIC/University of Seville, Seville, Spain, [4]Department of Computer Engineering, Automatics and Robotics, University of Granada, Granada, Spain

Synthetic data generated using generative models trained on real clinical data offers a promising solution to privacy concerns in health research. However, many efforts are limited by small or demographically narrow training datasets, reducing the generalizability of the synthetic data. To address this, we used real-world clinical data from nearly one million individuals with diabetes in the Andalusian Population Health Database (BPS) to generate a comprehensive longitudinal synthetic dataset. We employed a dual adversarial autoencoder to produce synthetic data and evaluated its utility in a clinical machine learning (ML) task: predicting the onset of chronic kidney disease, a common diabetes complication. Models trained on synthetic data were assessed for their ability to reproduce patterns and predictive behaviors observed in real data. Performance and stability were compared across models trained on real, synthetic, and hybrid datasets. Models trained exclusively on synthetic data achieved AUROC scores comparable to real-data models (0.70 vs. 0.73) and showed high stability in feature importance rankings (weighted Kendall's $\tau > 0.9$). Notably, combining synthetic and real data did not improve performance. Our findings demonstrate that high-fidelity synthetic longitudinal data can replicate real data performance in clinical ML, supporting its use in research while preserving patient privacy. This represents a significant step toward more collaborative and privacy-preserving healthcare data ecosystems.

SB6-03

Precision Prediction of Hepatitis C Using Population Health Data and Machine Learning

Alberto Esteban Medina, Isidoro Gutiérrez-Álvarez, Víctor Manuel de la Oliva Roque, Laura Alejos, Dolores Muñoyerro-Muñiz, Román Villegas, Joaquin Dopazo, and Carlos Loucera

Background. Early identification of individuals at risk of hepatitis C (HCV) within routine healthcare data could enable targeted testing and earlier linkage to care. Objective. To establish a transparent, temporally validated baseline for HCV risk prediction using population-scale electronic health records (EHR) from Andalusia and to outline a rigorous roadmap for improvement. Data and Methods. We constructed cohorts from the Andalusian Population Health Database (BPS) using routine laboratory results and diagnoses prior to an index date. Gradient-boosted decision trees (XGBoost) were trained on 2018–2021 data and temporally validated on an unseen 2022 cohort. We defined a prespecified evaluation suite including precision/PPV, recall/sensitivity, specificity, NPV, ROC-AUC, PR-AUC, calibration, and subgroup analyses. Results. On 2022 temporal validation, the current baseline model achieved precision ≈ 0.04 at the selected operating point under strong class imbalance (estimated prevalence ≈ 0.7%). Additional metrics (recall, specificity, NPV, ROC-AUC, PR-AUC) and error analyses are prescribed in this revision but require a full re-run to report numerically. Conclusions. The present work should be read as a baseline and methodology paper: results are modest but reproducible, leakage-aware, and temporally validated. We provide a complete pipeline description, stronger governance/ethics reporting, pre-registered metrics, and clinically motivated baselines to enable fair comparison in subsequent iterations.

SB6-04

Predicting Frailty in Older Adults Using Analytical Biomarkers and Machine Learning

Leasly Garcia Arana[1]
[1]*EAP Dreta de l'Eixample, Geriatric Department, Barcelona, Spain.*

Frailty is a multidimensional condition associated with increased morbidity and mortality in older adults. The VIG-Frail index, based on the deficit accumulation model and ranging from 0 to 1, is the standard tool for frailty assessment in Catalonia, Spain, but its routine application is resource-intensive. This study protocol aims to correlate the VIGFrail score with routine blood test parameters and to develop a predictive model for frailty using machine learning. A retrospective cohort of institutionalized and community-dwelling patients aged ≥80 years, or younger individuals classified as PCC or MACA, will be analyzed at the Primary Care Center EAP Dreta de l'Eixample. Eligible patients must have a valid VIG-Frail score and a blood test within ±3 months of the assessment. Laboratory values will be extracted automatically from clinical reports, while demographic and clinical variables will be entered manually. Gradient Boosting Regressor will serve as the primary predictive model, with Random Forest as a comparator, evaluated through 5-fold cross-validation and performance metrics (RMSE, MAE, R^2, and derived classification accuracy, macro-F1, AUC). SHAP will be applied to assess biomarker importance. The expected outcome could be a validated predictive model of frailty.

SB6-05

Early Risk Assessment for Ovarian Cancer Using Large-Scale Real-World Health Data

Víctor de la Oliva[1], Isidoro Gutiérrez-Álvarez[1], Alberto Esteban-Medina[1], Laura Alejos[2], Dolores Muñoyerro-Muñiz[2], Román Villegas[2], Joaquín Dopazo[1,3], and Carlos Loucera[1]
[1] Andalusian Platform for Computational Medicine, Andalusian Public Foundation Progress and Health-FPS, Seville, Spain, [2]Subdirección Técnica Asesora de Gestión de la Información. Servicio Andaluz de Salud. Sevilla, Spain, [3]Institute of Biomedicine of Seville, IBiS, University Hospital Virgen del Roco / CSIC / University of Sevilla, 41013 Sevilla, Spain

This research presents a novel machine learning approach for early identification of the risk of high-grade serous ovarian cancer (HGSOC) using large-scale real-world data from the Andalusian Health Population Database (BPS), which covers over 15 million patient records. We implemented an Explainable Boosting Machine (EBM) algorithm that processes heterogeneous clinical features including demographic information, chronic disease history, symptom profiles, laboratory results, and healthcare utilization metrics, eliminating the need for specialized biomarkers or predefined risk stratification protocols. Data preprocessing included conversion to the Observational Medical Outcomes Partnership (OMOP) common data model, ensuring algorithm portability across OMOP-compliant healthcare databases. The model development used 3,088 HGSOC patients diagnosed between 2018-2022 and 114,942 control subjects with comparable characteristics, maintaining an epidemiologically representative prevalence of the disease. Performance evaluation yielded 0.65 sensitivity and 0.85 specificity. This work demonstrates the effectiveness of population-level healthcare data for cancer risk prediction, offering a scalable and efficient alternative to conventional screening methodologies. The interpretable nature of our model provides transparency in risk factor identification, supports clinical decision-making, and potentially enables earlier therapeutic interventions with improved patient outcomes.