Abstracts of Seminars

Summer School on AI and ML in BME

25-28/6/2024

University camping of Kalandra, Chalkidiki

organized by the MSc in BME, AUTh
Biomedical signal processing and machine learning

Frantzidis, Christos

School of Computer Science, University of Lincoln, UK

Email: cfrantzidis@lincoln.ac.uk

This lecture is divided into three sections.

In the first section, students will learn about neurophysiological data acquisition during an object noun naming task. They will become familiar with the main steps involved in pre-processing biomedical signals to prepare the data for subsequent analysis.

In the second section, students will learn how to extract meaningful temporal information (amplitude and latency) from event-related potentials. They will use these features to train traditional machine learning classifiers, such as logistic regression, random forests, and support vector machines. Additionally, they will gain insights into classifier parameters, hyperparameter tuning, and training ensemble methods.

In the final section, students will apply their previous knowledge to the context of big data and deep learning approaches, specifically Convolutional Neural Networks. They will tackle real-world challenges, such as imbalanced datasets, and implement common mitigation strategies to achieve more meaningful results.
Hands on session: How to use Pytorch to segment lesions in mammography images.

Jakovljevic, Niksa
Faculty of Technical Sciences, University of Novi Sad
Email: jakovnik@uns.ac.rs

In this session students will find
- Why we prefer PyTorch over other similar tools
- What is a tensor and what we can do with it
- How to load DICOM files into tensors
- How we combine different data sources
- What are the basic building blocks in PyTorch
- How to use pretrained models
- How we trained our models for lesion segmentation
Novel ultrasound imaging markers of vascular function for prediction of cardiovascular events

Golemati, Spyretta

Medical School, National and Kapodistrian University of Athens

Email: sgolemati@med.uoa.gr

Vascular function is commonly assessed with ultrasound imaging; however targeted patient stratification remains a major clinical challenge. Coupled with advanced image analysis methods, conventional ultrasound imaging can yield novel markers of vascular function and thus address such challenge. Image-based morphological, textural, and mechanical indices can describe the complex physiological phenomena taking place in arteries and veins. Morphology and texture describe different patterns of tissue allocation, presumably as the result of exerted stresses, while mechanical features characterise tissue elasticity and are more sensitive to early tissue changes due to ageing or disease. Going one step further, the combination of such markers with machine learning can predict and justify adverse cardiovascular events, towards an in-depth understanding of vascular physiology and pathophysiology and a personalised patient stratification.
Biology-based Interpretation of Omics Data

Karamanidou, Dora

Pfizer Center for Digital Innovation

Email: Theodora.Karamanidou@pfizer.com

The presentation will focus on a bioinformatics platform for the interpretation of big biological data of the human genome. Through this platform, the user can attain more accurate and efficient genetic diagnostics and representative animal-to-human models to support clinical trials. The identification of causal genes can lead to better and faster diagnosis of various severe diseases, whereas the discovery of a more representative disease-to-animal model will add confidence to the translatable of disease processes across species.
To ChatGPT or not to ChatGPT: The role of AI and ML in Healthcare Education

Konstantinidis, Stathis

Health E-Learning and Media Team, University of Nottingham, UK

Email: Stathis.Konstantinidis@nottingham.ac.uk

While Artificial Intelligence and Machine Learning are not new, it is only recently that the research community put them in the forefront of research and practical application in all areas including Healthcare Education. This session will introduce the use of Machine Learning in education drawing examples from learning analytics, as well as the use of AI in education including examples from AI informed Chatbots and Large Language Models, Images development, article reviews and others, both from the side of the learner and the tutor. The second interactive part will allow the participants to use a free version of ChatGPT, and under guidance, they will develop a serious game: “Doctor for an hour”. The aim of the game will be to give the right medication to the virtual patient based on her symptoms, and cause and effect of the medication. There is no need for health/medical knowledge, as the medications, effects, causes and reactions will be imaginary and joyful. The aim of this game is to enhance the critical thinking of the players, but also to showcase to the participants how the AI LLM can be restricted to take the role of specific persona.
Applications and Use Cases in medical image analysis at the University Hospital Cologne

Li, Feifei

Biomedical Informatics Institute, University Hospital Cologne

Email: feifei.li@uk-koeln.de

In today's medical domain, machine learning methods, particularly supervised learning, face significant challenges due to distribution shifts in training and testing datasets, which often violate the i.i.d. (independent and identically distributed) assumption. These shifts can severely impact the robustness, fairness, and trustworthiness of medical image analysis applications. Addressing this issue, our research in UKK will explore the innovative use of CycleGAN (Generative Adversarial Networks) to cycle train CT (Computer Tomography) data from various scanners and manufacturers. This approach aims to mitigate distribution shifts by standardizing diverse data sources. However, the GAN-based model can produce images with artifacts due to model collapse and generative mechanisms. To refine these images and remove artifacts, we employ score-based diffusion generative models, enhancing the quality of the generated data while preserving critical features. Our evaluation, using five paired patients' medical images and SSIM (structural similarity index measure) metrics, demonstrates that CycleGAN serves effectively as a data augmentation technique but falls short in eliminating distribution shifts. Conversely, the denoising diffusion model proves more adept at addressing these shifts. Despite these advancements, the challenge of acquiring large, diverse datasets that reflect biological complexity remains. Future work will extend these generative methods under a federated learning architecture, aiming to broaden their application and effectiveness in overcoming distribution shifts across various medical imaging datasets.
OMOP-CDM, procedures of Patient Level Prediction or/and Population Level Estimation

*Natsiavas, Pantelis*

eHealth Lab, Institute of Applied Biosciences, Centre for Research and Technology Hellas

Email: pnatsiavas@certh.gr

AI is a buzzword nowadays. However, AI needs data. Focusing on health, data access is far from trivial due to the sensitive nature of health information and the need for privacy. The Observational Health Data Sciences and Informatics (OHDSI) ecosystem is perhaps the most prominent ecosystem supporting the real-world data analysis in healthcare for various use cases. The concepts of Patient Level Prediction and Population Level Estimation will be introduced.
Neuromarketing, a field standing at the crossroad of neuroscience and consumer behavior studies, provides an invaluable asset towards understanding decision making. Recently, several physiological monitoring modalities of complementary nature are jointly employed to recover a more complete picture. In this presentation, we will describe how we have used eye tracking together with electroencephalography so as to advance our understanding of the mechanisms underlying consumers’ decision making under a realistic scenario. Our study reveals that the integration of both modalities holds the potential of obtaining more accurate neuromarketing insights.
AI Applications in Impact Biomechanics for Crash Analysis

Oikonomou, Maria

Mechanical Engineer and PhD Candidate in BETA CAE Systems & AUTH

Email: maria.oikonomou@beta-cae.com

One of the automotive industry's main goals is to design an advanced protection system ensuring safe and convenient travel for humans. To achieve this, numerous virtual crash tests, known as crash simulations, are conducted to evaluate the safety of various road users in multiple crash scenarios. Continuous technological progress and emerging challenges, such as the introduction of Highly Automated Vehicles allowing diverse seating choices and the need to protect vulnerable road users like cyclists, the elderly, and the obese, necessitate innovative crash simulation tools and an in-depth investigation of human behavior in crashes. Currently, Crash Test Dummies and Human Body Models are the most common representations of humans in crash simulations, each offering distinct advantages.

This presentation aims to familiarize the audience with the applications of impact biomechanics in crash analysis, demonstrate the current challenges, and present the pioneering research of BETA CAE in the field. BETA CAE, besides developing practical tools that enable the biofidelic positioning of human models in a wide spectrum of impact scenarios, is also advancing the field with Artificial Intelligence applications. Machine Learning algorithms have been created to predict pedestrian head injuries in various crash situations without the need for corresponding crash simulations. Moreover, stochastic models are being developed based on experimental data to define representative cyclist postures for a wide spectrum of cyclists, aiming to conduct crash simulations to investigate their safety level.
Introduction to machine learning using Python

Papapanagiotou, Vasileios

Department of Biosciences and Nutrition, Karolinska Institute

Email: vasileios.papapanagiotou@ki.se

The presentation will introduce fundamental concepts and techniques, along with examples the Python programming language. Attendees will explore the basics of supervised and unsupervised learning, delve into key algorithms such as linear regression, decision trees, and clustering, and understand the importance of data preprocessing and model evaluation. Through practical examples and hands-on coding sessions, participants will gain the skills needed to implement and experiment with machine learning models using popular Python libraries such as Scikit-learn and Pandas.
Deep Learning for Accelerating Biomedical Knowledge Discovery

Passalis, Nikos

School of Informatics, Aristotle University of Thessaloniki

Email: passalis@csd.auth.gr

AI has witnessed exponential growth due to the development of powerful Deep Learning (DL)-based models that have enabled a wide range of impressive applications, ranging from autonomous cars to chatbots with immense knowledge and reasoning capabilities. Apart from these applications, DL also holds the potential for accelerating knowledge discovery in domains where research is bottlenecked by time-consuming and/or expensive experiments, often involving trial-and-error approaches. Last year, DeepMind's DL-based GnoME tool was successfully employed for discovering over 2.2 million novel materials, equivalent to almost 800 years' worth of knowledge. Similar findings have fueled research on the applications of DL in several fields of biomedical engineering, such as accelerating drug discovery and providing personalized medicine solutions. In this talk, we will provide an overview of DL approaches that can be employed for designing and optimizing experiments, offering a structured way to accelerate knowledge discovery in different fields, mainly focusing on expensive-to-evaluate black box functions/processes. We will cover aspects ranging from predictive modeling and active learning to Bayesian optimization and deep reinforcement learning approaches. Furthermore, we will also briefly discuss hybrid approaches, enabling the incorporation of prior domain knowledge into DL models, as well as the impact of model explainability and trustworthiness as a critical aspect of knowledge discovery. Finally, a use case on using an AI approach for probiotic strain selection will be presented, along with indicative hands-on examples.
AI in Medical Imaging

*Pitas, Ioannis*

School of Informatics, Aristotle University of Thessaloniki

Email: pitas@csd.auth.gr

This lecture overviews the relation between AI and medical imaging and diagnosis. First, an informative summary of “What is AI?” is presented, containing topics such as Symbolic AI, Data, Machine Learning (Clustering, Classification and Neural Networks). Topics that are related to book content creation, e.g., image processing, computer vision and natural language processing, are presented. The use of AI in proteomics (protein folding prediction) is also overviewed. The role of AI in medical diagnosis and education is also detailed. Finally, the creation of a medical data market is presented.

Respiratory volume measurements.

Bibliography

Introduction to ML using R

Psomopoulos, Fotis

Centre for Research and Technology Hellas (CERTH) Institute of Applied Biosciences (INAB)

Email: fpsom@certh.gr

With the rise in high-throughput sequencing technologies, the volume of omics data has grown exponentially in recent times and a major issue is to mine useful knowledge from these data which are also heterogeneous in nature. Machine learning (ML) is a discipline in which computers perform automated learning without being programmed explicitly and assist humans to make sense of large and complex data sets. The analysis of complex high-volume data is not trivial and classical tools cannot be used to explore their full potential. Machine learning can thus be very useful in mining large omics datasets to uncover new insights that can advance the field of bioinformatics. During this 3-hour practical session, the participants will be introduced to the machine learning taxonomy and the applications of common machine learning algorithms to health-related data. We will cover the common methods being used to analyse different omics data sets by providing a practical context through the use of basic but widely used R libraries. The session will comprise a number of hands-on exercises and challenges where the participants will acquire a first understanding of the standard ML processes, as well as the practical skills in applying them on familiar problems and publicly available real-world data sets.
Machine learning applications using OMOP-CDM

Rekkas, Alexandros

eHealth Lab, Institute of Applied Biosciences, Centre for Research and Technology Hellas

Email: arekkas@certh.gr

With the rise of real-world data and the strive for standardization of healthcare databases, one of the objectives of the global OHDSI community is the deployment of standardized methodological frameworks, aiming to improve health outcomes. We will demonstrate the PatientLevelPrediction R-package, developed and maintained by the OHDSI community used for the development and evaluation of state-of-the-art prediction models across a constantly expanding network of healthcare databases.

Turukalo, Tatjana Loncar

Faculty of Technical Sciences, University of Novi Sad

Email: turukalo@uns.ac.rs

This lecture will briefly introduce the basics of convolutional neural networks emphasizing their advantages in image analysis. Lecture will demonstrate how the convolutional layers are built, what is the number of parameters and what are hyperparameters in a CNN architecture. We will review advancements in architectures for image classification and present how to tackle the problem of semantic segmentation using CNN. The case study of suspicious lesion segmentation in mammography will be presented, illustrating our experiences from INCISIVE AI for health imaging project. We will talk about the challenges and approaches for harmonization of image sources, approaches to evaluate the performance and analysis of the results.