Proceedings of the 6th International workshop on knowledge discovery from healthcare data (KDH@IJCAI 2023).

IBRAHIM, Z., WU, H. and WIRATUNGA, N.

2023

Preface: The 6th International Workshop on Knowledge Discovery in Healthcare Data (KDH)

Introduction

The translation of routinely collected medical data into knowledge that drives the continual improvement of medical care poses grand technical challenges: 1) the extraction, organisation and assembly of large amounts of structured and free-text data embedded with electronic patient records, 2) near real-time analytics and knowledge discovery from the large, temporal, unevenly-sampled and uncertainty-ridden healthcare data and 3) overcoming data biases, problem and domain heterogeneity to design trustworthy prognostic and decision support models that are robust and align with clinical guidelines and workflow. Therefore, the successful design and implementation of tools that convert the data generated as a by-product of patient care into useful insight that can improve efficiency encompasses research in prominent areas of Artificial Intelligence including language engineering, data mining, knowledge representation and reasoning, learning and autonomous systems. This workshop is centred around novel AI methodologies that aim to solve some of the grand challenges associated with medical data.

Held in conjunction with the International Joint Conference on Artificial Intelligence (IJCAI 2023), this year's workshop continues the successful KDH series from 2016 to 2022 and builds on this year's IJCAI theme of 'AI for Good.' The workshop received 20 submissions, all of which underwent peer review by members of our program committee. Following the review phase, 5 long papers and 2 short papers were selected for presentation during the workshop.

Within the accepted papers, the workshop addresses topics in AI ethics, robustness, interpretability, and fairness. These themes are covered across various domains, including disease-specific prognosis, natural language processing, data mining and time-series analysis.

Invited Speakers

Dr Neil Hurley

Biography: Dr Neil Hurley is an Associate Professor in Computer Science, the head of the School of Computer Science at University College Dublin and a principal investigator at the Insight SFI Centre for Data Analytics, Ireland. His research spans data analytics, social network analysis, recommender systems, data hiding, digital watermarking, fingerprinting and high-performance computing. He has won over €1 million euro in research funding from Enterprise Ireland, Science Foundation Ireland, the European Union and industrial partners.

Talk Title: Health Informatics at the Insight Centre for Data Analytics

Dr Asif Ekbal

Biography: Dr. Asif Ekbal is an Associate Professor in the Department of Computer Science and Engineering at the India Institute of Technology (IIT) Patna, India. His research interests over the past 20 years are in Natural Language Processing (NLP), Information Extraction, Text Mining and Machine Learning (ML). He has authored over 300 papers in top-tier AI conferences and journals. Asif has been involved in several sponsored research projects, funded by different private agencies, such as Elsevier, Accenture, ezDI, LG, Skymap, Samsung Research, Wipro, Flipkart; and Govt. agencies such as MeiTY, Govt. of India, MHRD, Govt. of India, and SERB, Govt. of India etc.

Talk Title: Harnessing the power of NLP for improved clinical decision-making and patient care

Dr Atsushi Suzuki

Biography: Atsushi Suzuki was conferred a Doctoral Degree from the University of Tokyo. During the PhD study, Atsushi also worked as a research fellow on the Research Fellowship for Young Scientists offered by the Japan Society for the Promotion of Science. After obtaining the degree, Atsushi worked as an Assistant Professor at the University of Greenwich, United Kingdom prior to joining King's College London. His research interests span applied mathematics for data science, including learning theory, information theory and optimization, data science models and applications of data science.

Title: Hyperbolic-space-based approach for hierarchical data: its potential applications for healthcare

Accepted Papers

The following full papers presenting original research works were accepted. In Natural Language Processing, Moscato et al. detail a data augmentation framework for named entity recognition and an associated refinement that allows the selection of the most informative examples in an augmented data pool by minimizing noise. Zhang and Roberts present a framework for entity recognition and relation extraction from biomedical text. The framework is a generative model that bypasses the issues that arise when considering the two tasks in a pipeline manner, which ignores the interactions between the tasks.

Two contributions were centred around improving algorithmic performance in biomedical settings. In time-series analysis, Qian, Ibrahim and Dobson build on the state-of-the-art deep learning imputation models via a transformer-based architecture that scales to domains where the data distribution is highly skewed. Ali, Chourasia and Patterson present an evaluation of the effectiveness of Anderson Acceleration in aiding the convergence of machine learning algorithms using bioinformatics settings.

In condition-specific contributions, Kok et al. present an end-to-end machine-learning pipeline for detecting changes in breathing patterns in COPD patients, while Wu et al develop a multi-headed Transformer-based framework to predict viral mutation in SARS-CoV-2.

Finally, our workshop addressed implementation issues, especially pertaining to the fairness and trust of AI systems in sensitive domains such as healthcare. Yogarajan et al. propose an AI-based healthcare framework that comprises a feedback loop that facilitates quality improvements via continuous input provision.

We very much appreciate the support of the workshop chairs, Hadi Hosseini (Penn State University, US) and Viviana Mascardi (the University of Genova, IT). We sincerely hope that the participants enjoyed this year's workshop program and that this collection of papers will inspire and encourage more AI-related research for and within medicine in the future.

Zina Ibrahim, Honghan Wu, Nirmalie Wiratunga
Macao, 2023

Organisation

Workshop Co-chairs

- Zina Ibrahim, King's College London (UK)
- Honghan Wu, University College London (UK)
- Nirmalie Wiratunga, Robert Gordon University, Aberdeen (UK)

Steering Committee

- Kerstin Bach, Norwegian University of Science and Technology (Norway)
- Sadid Hasan, Philips Research North America (USA)
- Zina Ibrahim, King's College London (UK)
- Jonathan Rubin, Philips Research North America (USA)
- Nirmalie Wiratunga, The Robert Gordon University (UK)
- Honghan Wu, University of Edinburgh (UK)

Program Committee

- Felix Dransfield, King's College London
- Lorena Escudero, University of Cambridge
- Holger Kunz, University College London
- Lei Lu, University of Oxford
- Yvonne Lu, University of Oxford
- Stewart Massie, Robert Gordon University
- Tristan Naumann, Microsoft
- Linglong Qian, King's College London
- Farah Shamout, New York University Abu Dhabi
- Paul Taylor, University College London
- Anjana Wijekoon, Robert Gordon University