

Keynote Lecture

Trustworthy RL for Healthcare

Finale DOSHI-VELEZ (Harvard University, USA)

In many health settings, we are asked to make recommendations about new treatment policies given only data from current practice. Compounded with the often low quality and high missingness present in health data, the health setting is a particularly challenging one for applying reinforcement learning - even though there are many potential benefits if we can use prior data to improve future treatments. In this talk, I will first describe some RL algorithms that we have developed for health settings, with applications to decision-making for HIV treatments and in the ICU. Next, I will discuss our work to validate our algorithms. Some of this work is statistical, improving methods for off-policy evaluation. However, statistical checks are necessary, but not sufficient: I will conclude by describing ways we have incorporated various kinds of transparency into our approaches so that human experts can also use their domain knowledge to identify potential concerns.

Session 1 (time: 10:15 – 11:05)

(Talk 1) Childhood Diabetes On-Track with Data Science and Technology

LIM Soo Ting (KK Women's and Children's Hospital)

There is increasing evidence from research studies that suggest patient-centered care with positive experience has a good relationship with clinical outcomes. Historical authoritarian approach of care delivery and education has evolved to a care model encompassing shared decision making with patients and caregivers. The 20th century era of healthcare providers embrace evidence-based practice supported by data science and technology. Engage, Educate and Empowerment by technology-enabled Self-Monitoring of Blood Glucose for enhancing Self-Efficacy in Adolescents With Type 1 Diabetes is one of the highlights of how technology feeds data science in patient care journey which could be used for machine learning and artificial intelligence of future diabetes care model. However, precision medicine with technology depends on user interface requiring appropriate selection of patients/caregivers with mental capacity and readiness. In diabetes service for children and adolescents, specific care approach is catered to the needs of individual patient and caregivers which are crafted from value-driven projects with important clinical and nursing implications, as well as policy contributions. The talk on 'Childhood Diabetes On-Track with Data Science & Technology' will illustrate findings used by the diabetes care team in proposing a 'algorithmic' model of care for the adolescents living with Type 1 Diabetes Mellitus.

(Talk 2) Interpretable Machine Learning-Based Scoring System for Clinical Decision Making

LIU Nan
Duke-NUS Medical School

Scoring systems are widely used in clinical settings for convenient assessment of risks that provide important evidence for decision-making. With the increased availability of data and analytical tools, there have been ongoing efforts to update existing scores and to devise new clinical risk scores for a wide range of applications. In addition, scoring systems are inherently interpretable models. Such interpretability is highly valued for supporting medical decision-making, where doctors can easily understand how models make predictions. This talk aims to introduce and demonstrate how interpretable machine learning is used to derive scoring systems for decision making.

Session 2 (time: 11:05 – 11:55)

(Talk 3) Harnessing the Power of AI to Battle Covid-19 and Other Public Healthcare AI Application Examples

GOH Han Leong
Integrated Health Information Systems

Senior Principal Specialist of IHiS' Data Analytics & AI-Data Science & AI, Dr Goh Han Leong will share learnings from use cases in the development of Community Acquired Pneumonia and COVID-19 AI Predictive Engine (CAPE), an AI HealthTech Tool That Predicts Pneumonia Severity in Patients to combat COVID-19. Dr Goh will also share how IHiS has been working in partnership with Singapore's public healthcare institutions to create an expanding portfolio of analytics and AI solutions across four broad areas: (i) early stage detection and prevention, (ii) clinical treatment targeting as well as targeting for related special health support programs, (iii) resource optimization within hospitals and polyclinics and across the entire public healthcare supply chain and service delivery network, and iv) administrative support.

(Talk 4) Big Data and AI for Adaptive and Individualized Management of Cardiometabolic Disease

Pavitra Krishnaswamy
Institute for Infocomm Research

Abstract is not available

Session 3 (time: 13:15 – 14:30)

(Talk 5)

Neural collective matrix factorization for integrated analysis of heterogeneous biomedical data

Vaibhav Rajan
National University of Singapore

In many biomedical studies, there arises the need to integrate data from multiple directly or indirectly related sources: different sources may provide complementary information and one source may compensate for noise, uncertainty, or missing values in another. Collective matrix factorization (CMF) and its variants are models designed to learn latent factors of each constituent entity in arbitrary collections of matrices. The latent factors are rich integrative representations that can be used as features in machine learning models. After outlining classical CMF methods, I'll describe their inadequacies with respect to modeling biomedical data. We address these limitations by designing the first fully neural model for CMF. We have evaluated our models for various applications such as prediction of cancer drug-targets, adverse drug events, and drug responses, using genomic data, electronic medical records and other sources. Our experiments illustrate the versatility and efficacy of CMF-based neural representation learning for seamless integration of heterogeneous data.

(Talk 6)

DeepST: A Novel Graph Self-Supervised Contrastive Learning Method for Spatially Informed Clustering, Integration, and Deconvolution of Spatial Transcriptomics

CHEN Jinmiao
Singapore Immunology Network

Advances in spatial transcriptomics technologies has enabled gene expression profiling of tissues while retaining the spatial context. To effectively exploit the data, spatially informed analysis tools are required for three key tasks, spatial clustering, multi-sample integration, and cell type deconvolution. Here, we present DeepST, a novel graph self-supervised contrastive learning method that incorporates spatial location information and gene expression profiles to accomplish all three tasks in a streamlined process and outperform existing methods in each task. DeepST combines graph neural networks with self-supervised contrastive learning to learn informative and discriminative spot representation by minimizing the embedding distance between spatially adjacent spots and vice versa. With DeepST, we achieved 10% higher clustering accuracy on multiple datasets than competing methods, and better delineated the fine-grained structures in tissues such as the brain and embryo. Moreover, DeepST is the only method that can jointly analyze multiple tissue slices in both vertical and horizontal integration while correcting for batch effects. Lastly, compared to other methods, DeepST's cell type deconvolution showed higher accuracy on simulated data and better captured spatial niches such as germinal centers of the lymph node in experimentally acquired data. We further demonstrated that DeepST can recover immune cell distribution in the different regions of breast tumor tissue and reveal spatial niches with exhausted tumor infiltrating T cells. Our results show that DeepST is widely applicable to a broad range of tissue types and technology platforms. It offers a streamlined, user friendly and computationally efficient tool for characterizing tissue complexity and gaining biological insights into spatial organization and cell-cell communication within tissues.

Session 4 (time: 14:45 – 16:00)

(Talk 7) Advances and challenges of using AI in digital pathology

LEE Hwee Kuan
Bioinformatics Institute, A*Star

The field of pathology had undergone big changes due to the development of high throughput digital slide scanners. Slide scanners created in the era of digital pathology. Since pathology information has been digitalized in a large scale, applications of AI on these digital slides comes naturally. Indeed there are huge efforts worldwide in the use of AI for digital pathology. However, major challenges remains and we have not seen widespread adoption of AI in routine clinical use. In this seminar, we will discuss several latest AI techniques developed for digital pathology as well as outstanding challenges in this field.

(Talk 8)

Intelligent Ocular Image Processing – Research Update of iMED Team 2022

LIU Jiang
SUSTech, China

Dr. Liu will talk about the modalities, methods, algorithms of the ocular imaging research. He will also highlight the iMED team (<https://faculty.sustech.edu.cn/liuj/>) research progress in the past one year in ocular image processing.

(Talk 9)

Communication-Efficient Secure Federated Statistical Tests for PRS Validation from Multiparty Homomorphic Encryption

JIN Chao
Institute for Infocomm Research

The power and robustness of statistical tests are strongly tied to the amount of data available for testing. However, much of the collected data today is siloed amongst various data owners due to privacy concerns, thus limiting the utility of the collected data. While frameworks for secure multiparty computation enable functions to be securely evaluated on federated datasets, they depend on protocols over secret shared data, which result in high communication costs even in the semi-honest setting. To handle these limitations, we present methods for securely evaluating statistical tests such the Welch's t-test using multiparty homomorphic encryption (MHE). We tested and evaluated our methods on the Polygenic Risk Score (PRS) validation use cases against real world datasets and found that our method for computing the Welch's t-test required 100x less communication than equivalent protocols implemented using secure multiparty computation (SMPC), resulting in up to 10x improvement in runtime. Moreover, we also present a novel protocol to perform a table lookup from a secret shared index and use it to build a hybrid protocol that switches between MHE and SMPC representations in order to calculate the p-value of the statistics efficiently. This hybrid protocol is 1.5x faster than equivalent protocols implemented using SMPC alone.