# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Director's address</td>
<td>3</td>
</tr>
<tr>
<td>Alumni Highlights</td>
<td>5</td>
</tr>
<tr>
<td>Faculty and Staff Highlights</td>
<td>7</td>
</tr>
<tr>
<td>New Faculty</td>
<td>13</td>
</tr>
<tr>
<td>Student Dissertations</td>
<td>23</td>
</tr>
<tr>
<td>Student Publications</td>
<td>28</td>
</tr>
<tr>
<td>Contact Us</td>
<td>33</td>
</tr>
</tbody>
</table>
Dear Colleagues,

Greetings from Pittsburgh! I cannot express enough gratitude to all the scientists, companies, and global institutions that have made possible for us the vaccinations that directly reduce morbidity and mortality associated with COVID-19, at the population level. I realized this year that we (humans) are both the problem and the solution! I cannot help but remember the wise words of a famous Persian poet, Jalaluddin Rumi, who said (maybe in another context):

“You are not a drop in the ocean, You are the entire ocean in a drop.”

Our ISP has completed 34 years and is growing stronger each year! We are fortunate that our field of multidisciplinary applied AI has begun to attract new faculty members. This year, we had four new faculty members who joined our ISP. Primary affiliations of these new faculty, all of whom are featured in this newsletter, includes the Departments of Biomedical Informatics (DBMI), Computer Science (CS), as well as Informatics and Networked Systems (DINS). We also send our best wishes to Professors Stephen Hirtle (SCI, Psychology), and Michael Wagner (DBMI), two ISP faculty who retired this year, but continue to be active in mentoring students through committee service.

ISP welcomes a new batch of 7 doctoral students beginning this Fall 2021, with an additional MS student who is pursuing a doctoral degree in the Department of Philosophy currently. The ISP is highly selective, receiving almost 100 applications each year, and we congratulate these students on choosing our ISP for their doctoral degree pursuits. They will be featured on our new ISP website, which will be made public this month. Each Fall, the new website will feature Alumni of the ISP with significant achievements recognized by peers during the previous year.

One of the highlights of Spring 2021 was the joint SCI-DBMI-ISP Faculty Retreat held after a decade-long hiatus. This was very well attended by 46 faculty members across various departments within the University of Pittsburgh’s School of Computing and Information, School of Medicine, Graduate School of Public and International Affairs, and the School of Law to
name a few. It was extremely energizing to learn of the vitality and ongoing research of University Professor Emeritus, and Distinguished AI researcher, Prof. Bruce G. Buchanan; who appeared as a surprise speaker at this online event.

This year, our ISP Alumni, Faculty, and Students have accomplished major feats as evidenced by the large numbers of grants and publications mentioned herein. This growing ISP community has been ably supported by our one and only staff member, Michele. The time has come for us to grow even faster as applied multidisciplinary AI is becoming ubiquitous and embedded pervasively in man and machines. I hope that you enjoy every aspect of this newsletter, and will forward it proudly to your colleagues and friends, so they have no FOMO!

Sincerely yours,

Vanathi Gopalakrishnan, PhD
Director of Intelligent Systems Program
ALUMNI HIGHLIGHTS
Dr. Cristina Conati, Professor of computer science at the university of British Columbia, was nominated for *Outstanding Scientific Contributions to Computing* by the Association for Computing Machinery (ACM) in December 2020.

The ACM Distinguished Member program recognized 64 ACM worldwide membership based on professional experience as well as significant achievements in the computing field.

To be nominated, a candidate must have at least 15 years of professional experience in the computing field, five years of professional ACM membership in the last 10 years, and have achieved a significant level of accomplishment, or made a significant impact in the field of computing, computer science and/or information technology.

Dr. Shaghayegh (Sherry) Sahebi, Assistant Professor at Department of Computer Science at the State University of New York at Albany, received a *National Science Foundation (NSF) Career Award*, for the project "CAREER: Time-Aware Multi-Objective Recommendation in Online Learning Environments." The project is designed to advance learning in online educational systems by assisting students in achieving personalized, efficient, and continuous learning. Dr. Sahebi’s research team will investigate a new generation of educational recommendation and personalization systems that are capable of suggesting learning materials of heterogeneous types to students, according to multiple learning and behavioral goals, in continuous time. The total award amount is $547,705 for the 5-year grant.

Dr. Ioannis Tsamardinos, Professor at the Computer Science Department of University of Crete, Greece, and CEO and co-founder of Gnosis Data Analysis (Gnosis DA) and Just Add Data Bio (JADBio), presented JADBio, an AutoML tool for biological data, at Pitt in November 2020.

Dr. Matthias Grabmair joined the Technical University of Munich’s Department of Informatics as a tenure-track professor in Legal Tech in January 2021. Before joining TUM, he worked as a Legal Data Scientist at the German legal informatics company SINC.

Dr. Fattaneh Jabbari started a post doctoral fellowship at the Department of Biomedical Informatics at the University of Pittsburgh immediately after her graduation in September 2020, with Dr. Sofia Triantafillou as her research advisor.

Dr. Jaromir Savelka is a computer science post doctoral fellow at the Carnegie Mellon University since September 2020. He is currently a researcher at the Technology for Effective and Efficient Learning (TEEL) Lab.
FACULTY AND STAFF HIGHLIGHTS
6 YEARS CANCER FREE!
The ISP administrator, Michele Thomas, has been cancer free for six years! We wish her a happy and healthy life!

Dr. Vanathi Gopalakrishnan was mentioned in a news article for her contribution to a recent article in the AI magazine authored by Dr. Ganesh Mani, an adjunct faculty member at the Carnegie Mellon University. The article "Artificial Intelligence's Grand Challenges: Past, Present and Future," discusses the challenges AI faces in health, wealth, and wisdom, laying out a guideline for young researchers and professionals working in the field of artificial intelligent and its applications in different domains.

Dr. Yu-Ru Lin's article on "Policy Flow: Interpreting Policy Diffusion in Context" received the Best Paper Award for 2020 in ACM Transactions on Interactive Intelligent System.

Dr. Xiaowei Jia received the USGS award “Process-Guided Machine Learning for Modeling Stream and Reservoir Networks”. He also won the Best Applied Data Science Paper Award in SIAM International Conference on Data Mining 2021.

Dr. Diane Litman was recognized as an Outstanding Reviewer at the Conference on Empirical Methods in Natural Language Processing (EMNLP) 2020.
Dr. Shyam Visweswaran was elected as a Fellow of the American Medical Informatics Association (FAMIA) to the class of 2021 which recognizes the application of informatics skills and knowledge in a professional setting.

He was also awarded the Medical Student Research Mentoring Merit Award at the University of Pittsburgh Medical School in 2021. This award is presented to a Longitudinal Research Project (LRP) mentor of a graduating Pitt Med student in recognition for outstanding mentoring over the course of the LRP.

Dr. Panayiotis (Takis) V. Benos, received an R01 grant from NHLBI-NIH to develop interpretable probabilistic graphical models for multi-modal datasets. The methods will be used to identify factors causally linked to the progression of chronic obstructive pulmonary disease (COPD).

His lab also received an R01 grant from NHLBI-NIH to develop machine learning methods to identify and study chronic obstructive pulmonary disease (COPD) subtypes.
Pitt Momentum Funds

In the 2019–2020 academic year, Provost and Senior Vice Chancellor Ann E. Cudd and Senior Vice Chancellor for Research Rob A. Rutenbar introduced the Pitt Momentum Funds initiative as a joint large-scale research development fund. The Pitt Momentum Funds includes three categories:

- **Teaming Grants**: Support the early stage planning and capacity building of large multidisciplinary projects;
- **Seeding Grants**: Provide up to $16,000 per year to support significant and innovative scholarship by individuals or small groups of faculty;
- **Scaling Grants**: Support the detailed project planning, gathering of proof-of-concept results, and reduction of technical risk so that teams can competitively pursue large, complex, extramural funding.

Pitt Momentum Teaming Grants

Drs. Kevin Ashley, Diane Litman, Daqing He, Rebecca Hwa, and James Anderson were awarded the Pitt Momentum Teaming Grant 2021 on project: "Center for Text Analytic Methods in Legal Studies."

Drs. Richard Boyce and Malihe Alikhani were awarded the Pitt Momentum Teaming Grant 2021 on project: "Reducing language obstacles that Deaf students face when developing scientific competencies."
Dr. Morgan R. Frank received the Pitt Momentum Seeding Grant for his research proposal on "Detailed unemployment and the resilience of urban workers."

Dr. Xiaowei Jia was awarded the Pitt Momentum Seeding Grant for the project titled as "Physics-guided Machine Learning for Scientific Knowledge Discovery."

Dr. Sera Linardi and Dr. Yu-Ru Lin received the Pitt Momentum Seeding Grant for their research proposal on "Community-centric Building of Digital Infrastructures Against Systemic Oppression: G2A police misconduct complaint support system."

Dr. Seong Jae Hwang, Assistant Professor in the Department of Computer Science, received Alzheimer’s Disease Research Center Developmental Project grant by Pitt ADRC for 2021-2022 on "A bias-resilient deep learning algorithm for robust white matter hyperintensity segmentation on Alzheimer’s disease data with confounding factors" in December 2020.
Dr. Adriana Kovashka, Diane Litman, Rebecca Hwa, Malihe Alikhani, Yu-Ru Lin, and Jeffrey Cohn were awarded the Pitt Cyber Accelerator Grant on the project "Influencing Detection through Multimodal Discourse and Vividness Analysis," to explore the methods and ideas for extracting or completing agendas, detecting concerns and emotions, and modeling different types of vividness.

Kevin Ashley and Diane Litman have received a National Science Foundation (NSF) award titled as "Using AI to Increase Fairness by Improving Access to Justice," for their research on fairness in artificial intelligence in collaboration with Amazon.

Dr. Adriana Kovashka received an NSF CAREER award for the project "Natural Narratives and Multimodal Context as Weak Supervision for Learning Object Categories" in February 2021.

Dr. Kovashka also received an NSF award on "Generalization with shape and context" in August 2020.

Drs. Erin Walker, Adriana Kovashka, and Diane Litman received an NSF award for their project "Designing Effective Dialogue, Gaze, and Gesture Behaviors in a Social Robot that Supports Collaborative Learning in Middle School Mathematics" in September 2020.
Research and Career

What are your current research interests?
What is your vision for the future of your research field?

My research agenda is to embrace the complexity of artificial intelligence (AI), the future of work, and the socio-economic consequences of technological change. This goal is inherently multi-disciplinary and collaborative as it builds broadly on advances in the fields of labor economics, sociology, computational social science, network science, data science, political science, and complex systems. Rapid advances in AI and automation technologies have the potential to significantly disrupt the nature of work and workers' social identity. While AI and automation can augment the productivity of some, they can replace the work done by others, and will likely transform almost all occupations at least to some degree. Therefore, researchers, everyday workers, and policy makers must understand how technology will transform the nature of work.

These alterations accumulate and diffuse throughout the entire economy and contribute to today's growing economic inequality and accompanying socio-economic and political issues. Thus, I believe the frontier of my research field is to complement the traditional studies of economics, sociology, and political science by offering a complex systems perspective. Specifically, how do specific elements of a worker's skill set or of a labor market impact economic resilience, career mobility, and optimism about the future? Drilling down to the interacting elements that shape our economy can improve interventions and bolster workers for the future of work.
How do you expect joining ISP would affect your research and career?

Joining ISP offers a tremendously opportunity to grow my research agenda through collaboration and access to a variety of methodologies and viewpoints. My past experience have benefited from diverse multidisciplinary perspectives including computer scientists, physicists, economists, demographers, and urban developers. ISP offers a new way for me to solicit feedback, learn new approaches, and create collaborations to solve key problems in my field.

Are there any recent professional achievements (e.g., grants, nominations, awards) that you would like to share with ISP members?

I was awarded a Pitt Momentum Fund grant to study the economic resilience of US cities during the COVID-19 recession.

Courses and Publications

Do you plan to offer any new graduate-level courses?

I am offering Network Science and Analysis (INFSCI 2125) in the Fall and a seminar on computational social science (INFSCI 3350) in the Spring. I expect both courses would be widely appealing to ISP students.

Do you have any new publications that you would like ISP students to know about?


Advisees

How many graduate students are you advising now?

I am advising two PhD students at this time, and I am hoping to grow my lab.

Do you have any messages for the ISP students who wish to join your research group?

Please reach out if you are motivated to become an expert researcher in computational social science who uses data, statistics, complexity, and creativity to address major societal questions:

✉️ mrfrank@pitt.edu
🌐 www.pitt.edu/~mrfrank

Leisure Activities

How do you spend your spare time?

I love to get outside. I do a lot of mountain and road biking, running, and hiking. You can frequently find me in the woods with dog.
Have you recently read a book or an article that would like to suggest to ISP students?

Jump Starting America by Jonathan Gruber and Simon Johnson.

Advice to ISP Students

What advice would you give to new ISP students?

Be willing to work on lots of different projects with lots of different collaborators. My graduate career included studies of sentiment on Twitter, labor economics, international cooperation, and methodologies for online surveys. You will learn new ways to approach future problems if you can escape the dogma of the field in which you were trained.

Do you have any suggestions for ISP students who are close to graduation?

Present your work widely to many types of audiences. You never know with whom your research will resonate. Willingness to apply techniques across domains and sharing this findings can create interesting opportunities you may not have expected. For example, I would not have guessed as an applied math masters student that I would now be presenting research at management conferences and policy venues.

Also, use your social network and the social network of your collaborators.
Research and Career

What are your current research interests? What is your vision for the future of your research field?

My research interest includes data mining and machine learning. In particular, a major component of my research is to integrate scientific knowledge into deep learning for modeling complex spatio-temporal data sets.

Given the huge success of machine learning in commercial applications, there is a surge of interest in using machine learning to advance scientific discovery. However, direction application of machine learning has often met with limited success in scientific problems due to the paucity of training data, the complexity of underlying processes, and the uncertainty in data and models. Hence, there is an urgent need to develop a new generation of machine learning models that can represent complex processes and generalize to different scenarios.

How do you expect joining ISP would affect your research and career?

ISP provides a good environment for inter-disciplinary research. I have found many students in ISP are interested in the utility of machine learning in diverse applications. I am also able to find collaborators in different disciplines.

Are there any recent professional achievements (e.g., grants, nominations, awards) that you would like to share with ISP members?

I received the USGS award “Process-Guided Machine Learning for Modeling Stream and Reservoir Networks”. This project aims to build physics-aware machine learning solutions for modeling the water temperature dynamics in streams in response...
to climate change and the impact of human infrastructures (e.g., water release from reservoirs).

Our previous work on this project has won the Best Applied Data Science Paper Award in SIAM International Conference on Data Mining 2021.

Recent Courses or Publications

Do you plan to offer any new graduate-level courses?

Yes, I will teach a new graduate course “CS2756: Principles of Data Mining”, in Spring 2022.

Do you have any new publications that you would like ISP students to know about?


Advisees

How many graduate students are you advising now?

I have three students now.

Do you have any messages for the ISP students who wish to join your research group?

I am always looking for self-motivated students who are interested in developing data mining and machine learning algorithms for real-world problems of great societal relevance.

Leisure Activities

How do you spend your spare time?

Cooking, and playing with my son.

Advice to ISP Students

What advice would you give to new ISP students?

ISP is a great problem where students can learn the basics of artificial intelligence and its utility in diverse disciplines. I would encourage new students to actively reach out to faculty members to participate in real projects of their interest. This would help them learn how to build the AI pipeline for a real problem and what is still needed to advance existing AI algorithms.

Do you have any suggestions for ISP students who are close to graduation?

I wish they all will have a bright future. Artificial Intelligence is a continuously evolving field. I hope that they always keep the curiosity and enthusiasm to follow up with the development of this field through constant learning.
Dr. Hatice Osmanbeyoglu
Assistant Professor in the Department of Biomedical Informatics

**Research and Career**

*What are your current research interests? What is your vision for the future of your research field?*

My research focuses on developing predictive machine learning methodologies for dissecting the complexity of cancer using high-dimensional multi-omics datasets. We are interacting with domain experts (clinicians and cancer biologists) to identify statistical and computational problems and interpreting the results of new algorithms towards unlocking high-impact discoveries. Our field is developing rapidly, and AI can give researchers fresh insights into treating cancer and other diseases.

*How do you expect joining ISP would affect your research and career?*

My group is working on both clinical/biological applications and algorithmic development. I am looking forward to working with ISP students on developing new methods in advancing healthcare and our understanding of disease.

**Recent Courses and Publications**

*Do you plan to offer any new graduate-level courses?*

Not this year. I’M planning to offer a course about machine learning for single-cell genomics data analysis next year.

*Do you have any new publications that you would like ISP students to know about?*

My google scholar page is up to date if people want to see what my lab has been up to lately. We are also working on several papers right now which will be available on our website (www.osmanbeyoglulab.com) soon.

---

**Models Systems**

<table>
<thead>
<tr>
<th>Single cell</th>
<th>In vitro models</th>
</tr>
</thead>
<tbody>
<tr>
<td>In vivo models</td>
<td>Human data</td>
</tr>
</tbody>
</table>

**Data types**

<table>
<thead>
<tr>
<th>Genomics</th>
<th>Epigenomics</th>
<th>Transcriptomics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteomics</td>
<td>Metabolomics</td>
<td>Pharmacogenomics</td>
</tr>
</tbody>
</table>

**Regulatory pathways**

- Focus of research program
- Pan-cancer tumor classification
- Intra-tumor heterogeneity

**Precision medicine**

- Targeted therapy
- Immunotherapy
**Advisees**

*How many graduate students are you advising now?*

I am advising two masters students and three undergraduate students now. My postdoc recently moved to an independent faculty position. Two postdocs and two PhD students will be joining in August.

*Do you have any messages for the ISP students who wish to join your research group?*

I have openings for new students in my lab who are interested in the intersection of AI, machine learning, biology, and medicine. Students who tend to be a good fit are self-motivated, collaborative, want to make difference in biomedical domain.

**Leisure Activities**

*How do you spend your spare time?*

I really enjoy reading and painting.

*Have you recently read a book or an article that would like to suggest to ISP students?*

I recently read 3 Vital Questions by David Emerald. This book shows multiple ways to rise above problems and focus on outcomes.

**Advice to ISP Students**

*What advice would you give to new ISP students?*

Being open gaining experience in different domains.

*Do you have any suggestions for ISP students who are close to graduation?*

My advice would be to not limit themselves to defined career paths based on current job market. If they follow their passion with persistence and determination, they can carve out their own unique paths.
Dr. Malihe Alikhani
Assistant Professor in the Department of Computer Science

Research and Career

What are your current research interests?

My primary research interests are in the fields of Natural Language Processing and Cognitive Science with broader interests in Computational Social Sciences. I work towards building human-level artificial intelligent entities that can work with humans on a broad range of problems using a wide variety of methods, knowledge, and learning techniques.

Recent Courses and Publications

Do you plan to offer any new graduate-level courses?

Graduate NLP, AI Ethics, Computational Modeling of Discourse and Grounded Communication.

Do you have any new publications that you would like ISP students to know about?

For publications please see my 2021 papers from my website (www.malihealikhani.com)

Advisees

How many graduate students are you advising now?

I am advising four students at this moment.

Leisure Activities

How do you spend your spare time?

I enjoy hiking, reading and listening to music.

Have you recently read a book or an article that would like to suggest to ISP students?

• Justice: What’s the Right Thing to Do?: This book was published in 2009 on political philosophy by Michael Sandel, a Professor of
Government Theory at Harvard University Law School.


- *Educated*: is a memoir by Tara Westover on overcoming her survivalist Mormon family in order to go to college, and the role of education in her life.

**Advice to ISP Students**

*What advice would you give to new ISP students?*

Build effective research networks! No matter where you are or what your research area is, learning to communicate your ideas with other researchers and networking is an important part of your Ph.D. Attending and participating in local and international events, such as seminars, workshops, conferences, and lab meetings are a great way to build and maintain professional social relationships.

*Do you have any suggestions for ISP students who are close to graduation?*

Celebrate your achievements. Don’t compare your project to other people's projects or results. Look beyond publications. A PhD is a steep learning curve. Having the right work ethics and managing expectations can help you keep moving forward successfully.
STUDENT DISSERTATIONS
Fattaneh Jabbari successfully defended her doctoral dissertation in August 2020:

Instance-Specific Causal Bayesian Network Structure Learning

Much of science consists of discovering and modeling causal relationships in nature. Causal knowledge provides insight into the mechanisms acting currently (e.g., the side-effects caused by a new medication) and the prediction of outcomes that will follow when actions are taken (e.g., the chance that a disease will be cured if a particular medication is taken). In the past 30 years, there has been tremendous progress in developing computational methods for discovering causal knowledge from observational data. Some of the most significant progress in causal discovery research has occurred using causal Bayesian networks (CBNs). A CBN is a probabilistic graphical model that includes nodes and edges. Each node corresponds to a domain variable and each edge (or arc) is interpreted as a causal relationship between a parent node (a cause) and a child node (an effect), relative to the other nodes in the network.

In this dissertation, I focus on two problems: (1) developing efficient CBN structure learning methods that learn CBNs in the presence of latent variables (i.e., unmeasured or hidden variables). Handling latent variables is important in causal discovery since it can induce dependencies that need to be distinguished from direct causation. (2) developing instance-specific CBN structure learning algorithms to learn a CBN that is specific to an instance (e.g., patient), both with and without latent variables. Learning instance-specific CBNs is important in many areas of science, especially the biomedical domain; however, it is an under-studied research problem. In this dissertation, I develop various novel instancespecific CBN structure learning methods and evaluate them using simulated and real-world data.

Committee:

- Dr. Gregory F. Cooper, Intelligent Systems Program, University of Pittsburgh
- Dr. Shyam Visweswaran, Intelligent Systems Program, University of Pittsburgh
- Dr. Xinghua Lu, Intelligent Systems Program, University of Pittsburgh
- Dr. Peter Spirtes, Department of Philosophy, Carnegie Mellon University
Predicting the category of lexical items (e.g., words and phrases) is fundamental to a wide range NLP applications, from sentiment analysis to automatic essay grading. However, predicting the category of lexical items is difficult because the roles of the words and phrases depending on the context in which they are used. The existing approaches highly rely on the availability of the training corpus containing class labels for each lexical item, which is not available in many problem domains. In this thesis, we study the relationship between the lexical item classification problems and the corresponding, often easier to learn, classification problem at the greater context level (e.g., sentences). We propose an alternative training framework for lexical item category prediction, called inference by reverse contribution, which is only informed by the class predictions of greater contexts. We show that by developing a transfer function, which can rewrite a greater context from one class to another, we may infer the correlated class labels of the relevant lexical items.

To study the generalizability of the proposed framework to problem domains with different data availability profiles, in case where there is a limited amount of training examples for greater context prediction, we explore the transferability and adaptability of our proposed training framework to the new domains; in case where there is not a corpus of labeled training sentences, we discuss how data augmentation may help to reconcile the training requirements of the transfer function. Moreover, we study how our proposed self-attending and multiattending transfer function models may facilitate making lexical item prediction for problem domains with annotation profiles that involve phrasal and multiple lexical items.

Committee:

- Dr. Rebecca Hwa (advisor), Professor, Department of Computer Science, School of Computing and Information, University of Pittsburgh
- Dr. Diane Litman, Professor, Department of Computer Science, School of Computing and Information, University of Pittsburgh
- Dr. Adriana Kovashka, Assistant Professor, Department of Computer Science, School of Computing and Information, University of Pittsburgh
- Dr. Yu-Ru Lin, Associate Professor, Department of Information Science, School of Computing and Information, University of Pittsburgh

Omid Kashefi successfully defended his dissertation proposal in February 2021:

**A Training Framework for Resource-Constrained Lexical Item Class Predictions**
Di-Yang Xue successfully defended his dissertation proposal in April 2021:

Exploring Data and Clinical Knowledge Driven Computer-Aided Sequential Diagnosis System in Emergency Department

Medical diagnosis is the process of determining the nature of a disease and distinguishing it from other similar diseases. A diagnosis error happens when a diagnosis is missed, inappropriately delayed, or is wrong. Diagnosis error accounts for the most severe patient harm, the largest fraction of claims, and highest total penalty payouts. One way to reduce diagnostic error is to use a computer-aided diagnostic (CAD) system to augment doctors’ diagnostic abilities.

CAD systems can be roughly divided into two types: knowledge-driven systems and data-driven systems. Knowledge-driven systems strongly depend on pre-defined clinical knowledge. The usage of these systems may be restricted by the system’s knowledge base; they can only handle diseases and symptoms in the defined knowledge base. Extending and updating these systems is often labor extensive. Data-driven systems learn diagnosis rules from data directly, they can extend to new diseases very easily. Data-driven systems have three main drawbacks: black-box problems, not clinically meaningful, and not clinically actionable.

Few studies have been conducted on how to combine medical knowledge and machine learning algorithms to make a machine learning model clinically meaningful and actionable. We propose to (1) develop a framework that can integrate predefined medical knowledge with disease patterns in electronic medical records for disease differential diagnosis; (2) develop an algorithm that generates medical decision trees that recommend diagnosing actions by considering clinical workflow, diagnosis accuracy, and misdiagnosis costs. We plan to evaluate the performance of the developed system in real-world heart disease diagnosis tasks and compare its performance with other multi-label decision trees. We will also exam the effects of medical knowledge and electronic medical records on diagnosis accuracy, misdiagnosis costs, clinical meaningfulness, and actionability.

Committee:

- Dr. Daqing He, Professor, Department of informatics and networked systems, School of Computing and Information
- Dr. Gregory F. Cooper, Professor, Department of Biomedical Informatics, School of Medicine
- Dr. Michael M. Wagner, Professor, Department of Biomedical Informatics, School of Medicine
- Dr. Adam N. Frisch, Assistant Professor, Department of Emergency Medicine, School of Medicine
Amin Tajgardooin successfully defended his dissertation proposal in February 2021:

**Predicting Performance Deterioration using Divergence Metrics in Clinical Predictive Models**

Clinical predictive models are vulnerable to degradation in performance due to changes in the distribution of the data (distribution divergence) at application time. Significant reductions in model performance can lead to suboptimal medical decisions and harm to patients. Distribution divergence in healthcare data can arise from changes in medical practice, patient demographics, equipment, and measurement standards. However, estimating model performance at application time is challenging when labels are not readily available, which is often the case in healthcare. One solution to this challenge is to develop unsupervised methods of measuring distribution divergence that are predictive of changes in performance of clinical models. In this article, we investigate the capability of divergence metrics that can be computed without labels in estimating model performance under conditions of distribution divergence.

In this dissertation, we focus on two main objectives. (1) Investigating the capability of divergence metrics in detecting drift in clinical models. In particular, we use various dataset shift scenarios in electronic health record (EHR) data to measure the correlation between divergence metrics and the performance of models. The results of our preliminary experiments with state-of-the-art divergence metrics demonstrate significant correlations between dataset shift and drift in clinical models. (2) Investigating the ability of divergence metrics in estimating the efficiency of unsupervised domain adaptation (UDA) methods. We will perform extensive evaluations as well as develop efficient divergence metrics to examine the relationship between divergence metrics and the performance of UDA methods. Our proposed divergence metrics include semi-supervised and ensemble metrics. Semi-supervised metrics leverage labeled test data to detect changes in the distribution of target variables. Ensemble metrics combine dataset divergence metrics with metrics that use information from the model’s probability scores; this combination contributes an additional signal from the model that is not available to standard divergence metrics.

Committee:

- Dr. Shyam Visweswaran (advisor), Associate Professor, Department of Biomedical Informatics, University of Pittsburgh
- Dr. Harry Hochheiser, Associate Professor, Department of Biomedical Informatics, University of Pittsburgh
- Dr. Vanathi Gopalakrishnan, Associate Professor, Department of Biomedical Informatics, University of Pittsburgh
- Dr. Jeremy Weiss, Assistant Professor of Health Informatics at Heinz College, Carnegie Mellon University
STUDENT PUBLICATIONS

Sanya Bathla Taneja

- Li Sun, Ke Yu, Kayhan Batmanghelich, Context Matters: Graph-based Self-supervised Representation Learning for Medical Images, Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)

Mahbaneh Eshaghzadeh Torbati


Ke Yu presenting remotely at AAAI 2021

Saba Dadsetan


Anthony Sicilia


Amin Tajgardoona

Please follow ISP social media accounts for recent news and updates

You can contact the ISP Newsletter editor via email at isp.pitt@gmail.com

ISP newsletters are available at: www.isp.pitt.edu/newsletter

Editor:

Amin Tajgardoohn
mot16@pitt.edu