VOL 4 ISSUE 3 | SUMMER 2024 NEWSLETTER

Institute for Data Science

Hai Phan Authors Book to Provide Insights into Trustworthy Federated Learning

Associate Professor Hai Phan in Ying Wu College of Computing's Department of Data Science, and a pioneer in Trustworthy AI, has recently co-authored a handbook that will guide anyone from expert to beginner who seeks to venture into the realms of trustworthy federated learning.

"Handbook of Trustworthy Federated Learning" is published by Springer Publishing, a leading source of health care books, textbooks, and medical journals for medical professionals, professors, and universities, it is co-authored by My T. Thai, research foundation professor in the department of Computer & Information Science & Engineering at the University of Florida, and Bhavani Thuraisingham, founders chair professor of computer science at the University of Texas at Dallas.

continued on page 2

OVERVIEW:

- Student & Faculty Updates
- Recent Awards
- Publications

Hai Phan Authors Book to Provide Insights into Trustworthy Federated Learning

written by: Michael Giorgio
link to Full article

Associate Professor Hai Phan in Ying Wu College of Computing's Department of Data Science, and a pioneer in Trustworthy AI, has recently co-authored a handbook that will guide anyone from expert to beginner who seeks to venture into the realms of trustworthy federated learning.

"Handbook of Trustworthy Federated Learning" is published by Springer Publishing, a leading source of health care books, textbooks and medical journals for medical professionals, professors and universities, is co-authored by My T. Thai, research foundation professor in the department of Computer & Information Science & Engineering at the University of Florida, and Bhavani Thuraisingham, founders chair professor of computer science at the University of Texas at Dallas.

The book aims to be a "one-stop, reliable resource, including curated surveys and expository contributions on federated learning." It covers a comprehensive range of topics that combine technical and non-technical fundamentals, applications and extensive details of various topics.

First introduced in 2016, federated learning allows devices to collaboratively learn a shared model while keeping raw data localized to protect data privacy. Federated learning can enable myriad applications in practice, from health care and crime detection to human sensing based on mobile signals, among many others, without the prohibitive cost of privacy, security and administrative procedures of centralized data integration.

However, as Phan remarked in an earlier article recognizing a $740,000 grant to develop a trustworthy FL model to address new standards in AI safety and security, "Technology has advanced to a non-coding model that enables the general public to use AI in their everyday lives. However, if concerns over security and privacy prevent this, such advances are meaningless."

Phan and his collaborators spent the last one and a half years writing what they anticipate will be a "steppingstone" to broaden the future direction of optimizing privacy and security through the discovery and enhancement of large language models, while also helping to influence a greener carbon footprint.

IDS 2024 SUMMER HIGH SCHOOL INTERNS

First Name: Sai
School: John P. Stevens High School (Edison, NJ)

My summer internship allowed me to work alongside exceptional people, and it was an experience that seamlessly exposed me to the practical applications of data science. I was able to learn about algorithm implementation and how to approach errors in this area of research, through which I developed essential problem-solving and analytical skills. Moreover, hearing the insights of the rest of the team and the collaborative aspect of the research provided me with lasting communicative and teamwork skills. Overall, the experience was one that was truly enlightening and reinforced my drive to learn more about data science in a really unique and engaging manner.

First Name: Srijith
School: Edison Academy Magnet School (Edison, NJ)

My summer internship was an exciting opportunity to explore the field of combinatorial optimization. Through participating in real research problems and leaderboards, I was able to get a glimpse into what being a researcher at the forefront of an open problem is like. It was very exhilarating trying to beat the current score on the leaderboard of the problem we were working on, pushing the boundaries of computation every step of the way!

First Name: Isabella
School: Montgomery High School (Skillman, NJ)

I really enjoyed my experience this summer doing data science research under Dr. Bader and his research team. I learned a lot about data science through the research project, being able to get firsthand experience to learn more about the field. I really appreciate being given this opportunity and I hope to carry the skills and knowledge I gained from this summer in my future endeavors.

Structural Analysis of Biomedical Ontologies Center (SABOC)

2024 Grace Hopper Celebration

James Geller and his 11 students were joined by NJIT Provost John Pelesko at GHC 24' in Philadelphia, PA. The Grace Hopper Celebration empowers and inspires the next generation of women and nonbinary technologists through workshops, panels, presentations and more.

AI, Relying on Hardware Support, Could Improve by Thinking for Itself

written by: Evan Koblentz
link to Full article

People keep finding novel uses for generative artificial intelligence, the latest being that it can learn to design specialized hardware to make itself work faster.

Generative AI applications such as large language models became mainstream when ChatGPT went viral in 2022, but they require copious, complicated hardware underneath their user-friendly skins, especially when asked to act on more than just interactive text.

“Specialized [hardware] accelerators are crucial for maximizing the potential of AI tools, but current design tools’ complexity and required hardware expertise hinder innovation,” explained Arnob Ghosh, assistant professor in New Jersey Institute of Technology’s Electrical and Computer Engineering department.

Ghosh, along with colleagues Shaahin Angizi and Abdallah Khreishah, had the meta-idea to tweak a large language model as their assistant. They thought of training it to learn the context of what’s needed when designing hardware acceleration, based on a user’s needs for accuracy, energy usage and speed.

“We are trying to provide the optimal context so that an LLM can generate the desired results. This is based on the idea that the LLM can indeed demonstrate in-context learning,” Ghosh said. “The challenging question is how can we do prompt optimization here. Some basic instructions might not work. The prompt must consist of some of the elements of the codes themselves so that we can provide the optimal context to the LLM."

Their ideas include providing some hand-crafted instructions for basic hardware designs so the LLM has a basis from which to extrapolate its own creations, fine-tuning the model’s parameters for specific tasks, and using Khreishah’s graphical neural network to simplify how much virtual thinking the model must perform.
The trio are each focusing on part of the problem. Ghosh is optimizing the prompts and writing code that lets a large language model think about how to develop circuits, Angizi is working on non-traditional computing architectures and Khreishah designs the representation learning, which refers to how AI decides the format for interpreting your commands.

NSF Grant to Develop AI-Powered Solar Eruption Forecasting System

written by: Jesse Jenkins
link to Full article

New Jersey Institute of Technology (NJIT) researchers harness artificial intelligence for unprecedented insights into conditions in the Sun's lower atmosphere driving some of the solar system's most powerful explosions, capable of disrupting critical infrastructure on Earth.

NJIT researchers have been awarded a $593,864 National Science Foundation grant to develop a new AI system for more quickly and accurately predicting when explosive space weather events on the Sun will strike, from solar flares to coronal mass ejections (CMEs).

The three-year project, led by Yan Xu at NJIT's Institute for Space Weather Sciences (ISWS) and Jason Wang at the university's Ying Wu College of Computing, will develop AI-powered space weather forecasting capabilities that could offer solar researchers a new window into the complex magnetic processes in regions of the Sun's atmosphere that trigger such eruptions, and to this point, have rarely been observed. According to the researchers, the new AI-powered forecasting system — called SolarDM — could boost early-warning detection of these eruptive events on Earth by days, while offering vital insights to the space weather science community as activity on our nearest star ramps up over the course of the current 11-year solar cycle, which began in 2019.

"Major solar eruptions are powered by magnetic processes taking place in the solar corona, where we've lacked critical data due to poor observation conditions and insufficient instruments," said Xu, the project's principal investigator and research professor at NJIT's Center for Solar-Terrestrial Research. "Observations of the atmospheric layer underneath are crucial to study 3D magnetic fields. SolarDM's data insights potentially give us a way to map the magnetic landscape of this region, allowing us to better predict these powerful eruptions."

Solar physicists have long studied the structure and evolution of magnetic fields in the corona (the Sun's upper atmosphere). The breaking and reconnecting of these field lines are known to power explosive events capable of disrupting technologies on Earth, such as satellite operations.

AI Visionaries Podcast: Exploring the Future of AI and HPC with Dr. David A. Bader

Podcast series, Hosted by ZINFI CEO and Founder Sugata Sanyal and sponsored by Databank, where we delve into artificial intelligence and high-performance computing features distinguished speakers like Dr. David A. Bader, who share their insights on cutting-edge technologies and their impact on various industries.

Tune in to explore groundbreaking research, future trends, and the transformative potential of AI and HPC in today's rapidly evolving technological landscape. Click Here to listen.

Publications

28th Annual IEEE High Performance Extreme Computing Conference

Garrett Gonzalez-Rivas, Zhihui Du, David Bader (2024). A Deployment Tool for Large Scale Graph Analytics Framework Arachne. IEEE HPEC 2024. Click Here.

Fernando Vera Buschmann, Zhihui Du, David Bader (2024). Enhanced Knowledge Graph Attention Networks for Efficient Graph Learning (Outstanding Student Paper Award). IEEE HPEC 2024. Click Here.

Mohammad Dindoost, Oliver Alvarado Rodriguez, Sounak Bagchi, Palina Pauliuchenka, Zhihui Du, David Bader (2024). VF2-PS: Parallel and Scalable Subgraph Monomorphism in Arachne. IEEE HPEC 2024. Click Here.

Student Spotlight

Oliver Alvarado Rodriguez

4th year Doctoral Student in Computer Science
Participated in an Internship this summer with Hewlett Packard Enterprise working on the Chapel programming language team to optimize distributed graph algorithms.
Will be joining HPE as a Software Engineer III working on the Chapel programming language development team that is a part of the High-Performance Computing Advanced Development Organization at HPE

INSTITUTE FOR DATA SCIENCE DIRECTOR'S OFFICE

DAVID BADER

Institute Director

david.bader@njit.edu

SELENNY FABRE

Business Manager

selenny.m.fabre@njit.edu

About Us | Contact Us | Subscribe