AI 4 Science Discovery Network+

AI4SD Interview with Yingfang Yuan
06/12/2021
Online Interview

Michelle Pauli
Michelle Pauli Ltd

12/08/2022

AI4SD Interview with Yingfang Yuan
Humans-of-AI4SD:Interview-37
12/08/2022
DOI: 10.5258/SOTON/AI3SD0236
Published by University of Southampton

Network: Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery
This Network+ is EPSRC Funded under Grant No: EP/S000356/1

Principal Investigator: *Professor Jeremy Frey*
Co-Investigator: *Professor Mahesan Niranjan*
Network+ Coordinator: *Dr Samantha Kanza*

Contents

1 Interview Details

Title	AI4SD Interview with Yingfang Yuan
Interviewer	MP: Michelle Pauli - MichellePauli Ltd
Interviewee	YY: Yingfang Yuan - Heriot-Watt University
Interview Location	Online Interview
Dates	06/12/2021

2 Biography

Figure 1: Yingfang Yuan

Yingfang Yuan: 'If you don't have any problems, you don't really have any research'

Yingfang (James) Yuan is a PhD candidate working under the supervision of Dr Wei Pang, Prof Mike Chantler and Prof George M. Coghill (external) in the School of Mathematical and Computer Sciences at Heriot-Watt University. Yingfang received his MSc in Big Data and High-Performance Computing at University of Liverpool. His research interests include machine learning; deep learning (especially in Graph Neural Networks and AutoML). He particularly focused on investigating the impact of hyperparameter optimisation on graph neural networks applied to predict molecular properties and the efficiency of hyperparameter optimisation approaches.

In this Humans of AI4SD interview he discusses the challenges involved in graph networks and hyperparameter optimisation, making the most of interdisciplinary research, how AI can speed things up, and his advice for early career researchers.

3 Interview

MP: What's been your path to where you are today?

YY: My diploma was in computer science, but on the side of software development, which taught me some fundamental programming skills. During my Master's in computer science, I was looking for more at algorithms and machine learning techniques in relation to big data and high-performance computing. It felt so powerful and helpful to be involved in developing machine techniques that were used to solve real-world problems.

After my graduation, I hadn't planned what I wanted to do, so I stayed foolish and stayed hungry! I wanted to keep learning, so I applied for a PhD at the University of Aberdeen. The PhD involved two aspects: the first was about the learning — deep learning and optimisation algorithms — and the second was about the application — ensuring they had real-world application. It was lucky that my PhD gave me the opportunity to work on joint projects, which involved people from different fields, like material science or project management. There, I could apply my research to solve joint problems, for example, predicting chemical reactions.

MP: What research are you working on currently?

YY: There are two parts to my research: graph networks and hyperparameter optimisation. The research involves combining the two together, applying hyperparameter optimisation and graph networks to make molecular property predictions. I chose to work on molecular property prediction because of its real-world applications, for example, in drug discovery. Drug development is often quite time-consuming, it can take around six years to find a drug. If we can build models of graph neural networks, experts will be able to use our model to facilitate their work.

MP: What kinds of challenges have you faced in this research?

YY: There have been quite a few! Deep learning is moving at a fast pace. If you have an idea and you Google it, you will probably find that it's been conducted by others: that's the first challenge.

Our research can also be quite challenging because the architecture of a neural network can be quite complex and complicated. When you find an error, or something unexpected, it can be hard to verify the reason behind it. There are multiple hyperparameters in neural networks, so you can't isolate one at a time. In addition, hyperparameter optimisation is always time-consuming, not easy.

MP: Are there any challenges around interdisciplinary research?

YY: Communication between different fields can often be a challenge, and there's also the issue that people from different fields tend to focus on different aspects of a problem. Faced with one problem, a computer scientist and a chemist would look to different parts in order to solve it. But it's just like climbing a mountain, some choose path A, and others choose path B. The best way to overcome this is to facilitate discussion, having people sat together, negotiating and cooperating.

MP: What has surprised you in your research?

YY: It was a surprise just how much time could be saved by training deep neural networks. If you want to try different hyperparameter settings, it can take up to a month, but to train the model our way only takes around a week. Our method discards the unpromising hyperparameters at a very early stage, so you can use the time more efficiently to look for promising ones.

I've also been surprised by the usefulness of deep learning and graph networks to molecular problems. There's a lot more research now exploring the potential here.

MP: How is AI / machine learning changing how we do science?

YY: The first thing that comes to mind is the rate at which it speeds things up. When we think about scientific discoveries in the past, these involved long-term processes. But with machine learning and deep learning, this issue is alleviated dramatically. In traditional computational chemistry, for example, people use density function theory (DFT), and DFT computation is very costly. Now, a lot of people are trying to use graph networks to do the same/similar work because of the time they can save.

MP: What advice would you give to early career researchers?

YY: I would suggest that PhD students find a topic or area they find interesting first, before trying to read all the papers and accumulate knowledge. I would also suggest not worrying too much about the potential problems: having a lot of problems is good because you will be able to think about how to solve them through your research. If you don't have any problems, you don't really have any research!