

## Diaries of a Young Indian Research Scholar

“It's not that I'm so smart, it's just that I stay with problems longer.”

Albert Einstein

It is a privilege to get this platform to share my experiences as a young researcher in India. The experience of a data science researcher in India has its own nuances and flavours. This blog post is not a do's and don'ts list for pursuing data science or AI research in India, I don't believe such a characterisation exists, but this is merely my perspective acquired during my time as a Ph.D student.

About my background, I recently submitted my Ph.D thesis titled “Symmetry Aware Inference and Decision Making in Probabilistic Models” at IIT Delhi and joined Deepmind, Montreal in November 2018. Prior to my Ph.D, I completed masters at Indian Institute of Science, Bangalore and my bachelors in engineering from Punjab Engineering College, Chandigarh. I joined my Ph.D at IIT Delhi in Algorithms but later switched to ML/AI under the guidance of Prof. Parag Singla and Prof. Mausam.

The global landscape of ML changed during my time as a PhD student and I was lucky to witness this as an active researcher instead of a mere spectator. When I moved to data science in 2014, I was unaware that a company named DeepMind had already been acquired by Google. DeepMind's core technique in AlphaGo -- Reinforcement learning (RL), went from being a sidebar in the first lecture in ML classes to an entire course in itself. The change in the field was driven by the joint efforts of academia and industry.

My research journey started with joining Data Analytics and Intelligence Research (DAIR) group at IIT Delhi. Starting slowly and steadily, our group soon caught pace with some top tier publications. A major benefit of participation in these conferences was the interaction and connection with leading figures of machine learning and artificial intelligence. I must say, the community of ML/AI is very open and welcoming, wherein they actively promote inclusivity in research throughout the world. To facilitate this, several students get to register free for otherwise expensive conferences in return for volunteering. Moreover, the travel grants of the conference itself, as well as the Industries like Microsoft Research India, Xerox Research, TCS research and IARCS, further make conference travel affordable for people living in developing countries. I am grateful to these agencies for sponsoring my travel and participation in these conferences. Overall, travelling to international conferences become a lot easier in countries like India with these multiple grants which complemented the support from parent institutions.

Overall, the Indian data science community witnessed phenomenal growth in last few years. Many of the IITs, IIITs and IISc students have become regular publishers at top tier conferences. But, at the same time I feel a lot more needs to be done to bring our research ecosystem closer to the research havens of the west. I believe the following steps could be helpful in making our ecosystem stronger.

i) **Inter institution communication and collaborations:** It is difficult to find the best of all areas at any one institution. For example, IIT Madras has a strong RL group, IIIT H has strong vision group while IIT Kanpur and IISc have strong theory and optimization groups. Inter institution collaboration can potentially take up ambitious projects and attack hard problems. A collaborative and communicative environment would create a comprehensive experience at par with anywhere in the world, for any data science student in India. This could be encouraged by joint courses in multiple institutes, exchange programs among students or more visits and short term courses. I would love to see a joint forum where flow among indian institutions is relatively easy than current trends.

ii) **Research Start Ups:** Although there has been a surge in product based startups, pure research startups are still mostly unheard of in India. I wish some of us utilise our expertise to create research institutions which use the huge amounts of indigenous data at our disposal to solve interesting research problems. This could vary from health care startups using AI for direct applications to solving some of the most challenging research problems.

iii) **Open Sourcing Datasets/Softwares and end-to-end research:** We used to discuss a question : What is the most popular library released completely by an Indian institution which is used throughout the world? Though we have some early answers, the research at many places in our institutes is limited to publishing a paper at top tier conferences. This trend should be augmented with the release of open source codes and datasets. I would like to see a strong effort towards datasets, libraries and source codes releases jointly by multiple indian institutes.

I was also asked to give some nuggets of advice(*gyaan*) for students starting a research career in data science in India. On this front, I would like to share some of the lessons learnt during my research journey or from some of my own mistakes. The foremost advice for any researcher in this field is to build a strong foundation through courses. A strong background in linear algebra, probability and statistics and optimization and algorithms is essential to advance research in data sciences. Don't hesitate in taking courses throughout your Ph.D (even fourth or fifth year) since this time is a great learning period which you will miss if you don't utilise it optimally during your doctoral studies. Secondly, do at least two internships abroad with one of them being an industrial internship. This will help you to bridge the gaps in research between India and world and also, to understand the working environment at different places and cultures. Thirdly, though travel is a personal choice, but don't miss any opportunity to travel and attend conferences, summer schools and various leading forums. Lastly, love doing the teaching assistantship work. Don't consider it a burden, since this is an opportunity to see how a teacher thinks and a chance to be an entrepreneur in creating a new course..

Ph.D provides an opportunity to explore some new things in these five years, so explore some unique experiences like trekking, sports , retreats, dance, music and or develop a new hobby. Fill the glass of your student life with as many experiences as you can :) .

Overall, I believe that the data science is an exciting field today with many new opportunities coming up in years to come. I wish India could play a central stage in this field by novel, genuine and adopting finest ethical and professional practices. We are a small research community and so it is easy to imbibe these practices in our foundations at this time which would help us to create a beautiful research ecosystem in our country in the years to come.