Autobiography

Hoang Le

I see myself as someone who's always learning, especially in research. I try to make complex models more understandable for my classmates in data analytics. My background includes minors in computer science and mathematics, focusing on algebra, to get a better grasp of the tools we use in data analytics. I've also explored statistics, probability, and machine learning through Andrew Ng's courses to broaden my skill set.

I'm really into the idea of automation and how it can handle routine tasks like classification or calculations. It's exciting to think that automation can free us up to tackle more creative and abstract challenges. Oddly enough, I get a kick out of the trial-and-error process in problem-solving, tweaking layers and algorithms, and finally getting a model to work just right. I'm passionate about creating open-source tools that add value to our community.

Currently, I'm focusing on regression modeling and starting to get into Time series analysis, particularly with sound and audio. I haven't taken an advanced math class in this yet, so it's still a bit of a challenge, especially finding the right datasets that are relevant and consistent. For this seminar, I've prepared a mix of statistical models and applied probability knowledge to give me more flexibility. I've done some publication writing in the past, so I feel confident in my ability to write clear and informative research papers.

This semester, with Dr. Sarah Supp's guidance, I'm eager to dive deeper into ecology. With a variety of tools at my disposal, I'm looking forward to improving my skills and learning from

your expertise. I would love to continue with Dr. Supp's recent publication. I'm thinking of exploring the key factors and their interactions that contribute to biodiversity. Maybe I'll use time series analysis to see how certain features and interactions vary over time. This could potentially lead to a model that applies not just to one species or area but to a broader problem.

I'm into machine learning, though I'm not quite at the level of an AI engineer yet. My ultimate goal is to lead a data science research team in an industry setting. In a year, I'll start as a business analyst, mainly applying straightforward statistical models to solve business problems. I'm planning to dive deeper into programming and linear algebra during this time.

After my three years of OPT, I aim to pursue a Master's degree in machine learning or AI. This should set me up for an entry-level role as a Machine Learning Engineer or AI Engineer. And, if time permits, I'd love to follow that up with a postdoc, which I believe would greatly enhance my chances of landing a role as a Junior or Senior Data Scientist or Engineer.

I've come to believe that studying any subject can be manageable. The real challenge often lies in identifying the core knowledge gaps that need to be filled to learn effectively and efficiently. This is especially true in fields like programming and linear algebra, which are crucial not just for data scientists but also for data analysts. Nowadays, the roles of data scientists and analysts are increasingly overlapping. In my view, a data analyst is akin to a technician. It's vital for them to thoroughly understand the tools they use, as this understanding is key to advancing in their career.