

Cover Letter for Chalmers PhD application

Attila Nagy

April 28, 2018

Introduction and Qualifications

My journey in the world of machine learning and artificial intelligence started during my Computer Science Master's program at the Gothenburg University when I decided to take Artificial Intelligence course. On this course, I had the chance to be part of a team of 3 who developed a solution for a sentiment analyse of movie reviews and another solution for SHRDLite, a simpler version of SHRDLU[2], planner using GF for naturel language processing. One of the supervisors for this course was Devdatt Dubhashi. After my graduation, I directly started to work on Volvo's self-driving car project as a Software-Developer in the Sensor Fusion team. The work mainly involved post-processing of data from various sensors, such as HERE HD maps, Velodyne LiDAR, and development on Nvidia's Drive PX platform. As for the algorithm part, my involvement was less significant, but still relevant enough to understand and follow the development of Particle-Filtering for mapping and localization. After my 2 years at Volvo Cars, I felt that I need change in my life, and therefore decided to try out the digital nomad life-style and travel in South America and Southeast Asia. During this period of my life, I finished Udacity's Deep Learning Nanodegree program, and gained practical knowledge by free-lancing on Codementor. Currently, I work for a startup, RMBLStrip, located at Stena Center, next to Chalmer's Johanneberg campus. The main product for this company is an adaptive roof-deflector for the heavy truck industry. My role involves several parts of the development process, including data-collection and SW engineering, but the most interesting is algorithm development. One of our solutions uses a classical Bayesian probabilistic method to find the optimal position for the roof-deflector.

Motivation and Carrier Goals

As for my motivation for becoming a PhD, in the recent years, starting with the Nanodegree on Udacity, my interest drifted more and more towards understanding the theoretics of machine learning. During these 4 months of study I had the opportunity to try and see many state-of-the-art algorithms and tools in action. One of such tools was Google's TensorFlow for implementing 1) a Convolutional Neural Network for image classification using the MNIST dataset, 2) script generation for a fictional conversation between Homer and Moe in Moe's tavern for The Simpsons TV series using a Recurrent Neural Network with word2vec embeddings, 3) language-translation from English to French using RNNs as encoder and decoder in an auto-encoder, and finally 4) image-generation using Generative Adversarial Networks introduced by Ian Goodfellow himself as a guest lecturer. Another essential tool was a GPU cluster necessary for training such algorithms. During these 4 assignments, I was first using Amazons Web Server, but due to unpredictable speed variations and connection breaks, I decided to use FloydHub instead. Their service proved to be much more reliable and better fit to my needs. However, finally I ended up using my own NVIDIA graphic card in my desktop PC due to its limitless accessibility and similar performance in terms of GPU speed. All of the above mentioned projects were more focusing on practicalities than understanding the theoretical background of the algorithms. Most of the blogs that explained the essence of these algorithms barely scratched the surface of the topic and did not give a profound understanding in it. To gain such a knowledge, it is inevitable to deep dive into mathematics and fiddle with research papers. For instance, all the above mentioned assignments involved the use of dropout to gain better regularization, but the course introduced barely the basic idea and neglected to go further in the details. My main motivation for seeking a PhD position is to have the chance to gain a deeper understanding in the underlying concepts of machine learning and to contribute to the next advancements in the field. In my current employment these drives are let not to flourish due to incompatibility with the corporate interest, compelling me to pursue my interest in machine learning further as a hobby in my free time, after work. Hence is my motivation for applying for a PhD position at Chalmers in the field of mathematics in machine learning and artificial intelligence.

Phd Positions

From the 6 open PhD positions in machine learning at Chalmers the following two drew my attention:

1. **Ref 20180233**: Understanding deep learning: From theory to algorithms
2. **Ref 20180221**: Learning with noisy labels

Understanding deep learning

Understanding deep learning from a mathematical perspective is one of my key interesting in a PhD, and therefore this position is very appealing to me. Furthermore, this is the position that is broad and flexible enough that I can absolute see myself working with in the next 5 years. Understanding the theoretical background in mathematics for a stochastic regularization method, such as dropout, would be an interesting approach to start in this position. Therefore, I feel the most inclined to work in this field from the currently available 6 open PhD positions on Chalmers.

Learning with noisy labels

Learning with noisy labels was the other topic that immediately drew my attention. This is a recurrent topic in machine learning, significantly affecting all the segments in the field including both academia and industry. Furthermore, in this position I would have the chance to further polish my knowledge in probabilistic modelling that I recently gained at RMBLStrip working with a Bayesian Linear Regression algorithm. However, I do not feel fully qualified for this position since I have never officially taken any course in Information Theory. All my knowledge in this field is learned on-demand during work or my Udacity Nanodegree program. My favourite source for learning probabilistic programming during this period was the Bayesian Methods for Hackers book[1], that is available as a jupyter notebook online as well, making the whole learning process more interactive. The focus in this book is to provide a practical knowledge as well as the theory by having a handful of hands-on exercises using the PyMC3 Python library.

Other positions

Other position in Ref 20180221

References to contact

References

- [1] C. Davidson-Pilon. *Bayesian Methods for Hackers*. In [1], first edition, 2015. <https://github.com/CamDavidsonPilon/Probabilistic-Programming-and-Bayesian-Methods-for-Hackers>.
- [2] T. Winograd. *Procedures as a Representation for Data in a Computer Program for Understanding Natural Language*. PhD thesis, Feb. 1971.