

Report of Internship at Toyota Technological Institute at Chicago

Yota Toyama (Student ID: 16423), Computational Intelligence Laboratory

Life

After I and Tsujimura got the O'Hare airport, we took a train and a bus to go to our apartment in the southside of the downtown. When getting off a bus, we were just standing to wait for its doors to be open. But, they did not. After a while, passengers behind us told that we should push the door. We had never known such semi-automatic system. People in Chicagoo were very kind and helpful.

Our apartment was pretty good. It was much larger than we thought. And, we could see great views of the Michigan lake and Chicago city. It also has its own restaurant and grocery store, which were convenient when we do not want to go outside at cold night.

There was a supermarket named Whole Foods near our apartment. It was a little more expensive than other ones but sold healthy and organic foods of good quality. I thought that foods in the US are much cheaper than in Japan. However, that was not true at least in Chicago. Some stuffs like meat seemed to be even more expensive than ones in Japan. It can be because Chicago is a big city and they can be cheaper in rural areas, I guess.

When I arrived there, I had hay fever which I never had in the same season in Japan. So, I needed to buy medicine against it. They were pretty expensive by 30 dollars per bottle.

We did not have any difficulty to get good foods around TTIC. There are many shops and restaurants around it. In an early week, some students at TTIC took us to a Asian noodle restaurant called Noodles. I ordered a Singapore noodle and it was really good. I also became relieved because it was confirmed that we can get some Asian foods there. In weekdays, we usually went to food trucks to get foods with our Chinese friends. They sell variety of them, such as hamburgers, sandwiches, tacos and rice bowls. A few weeks after classes started, Chinese food delivery by a restaurant, Sweet Station also started. It also dealt in many kinds of Chinese foods from lunchbox-style ones to noodles.

Study

Class

I took a class of Machine Learning, titled Introduction to Statistical Machine Learning. At first, I was also taking a class of Mathematical Toolkit. But, I

found that it took too much time from me so that I cannot have time to work on my research. Therefore, I canceled it and focused on the ML class and my research task.

In the ML class, we studied about themes of wide range from basic ML to its application and some state-of-the-art stuffs like Neural Network. The models used in the class were supervised ones mostly. In 3 to 4 of the firsts classes, we learned about base of ML of probability, algebras, model complexity, overfitting, and bias and variance.

Its assignments were exciting. As in Japan, we had some mathematical problems of symbolic calculation and proofs. In addition to them, we had competitions in our homework a few times. To hold the competitions, the professor of the ML class used an online service named Kaggle. It is famous among ML and data science researchers and students and holds some public competitions of ML and data analytics. We had 4 competitions of multivariate logistic regression of MNIST hand-written digit classification, linear and non-linear SVM of sentiment analysis, Gaussian mixture model, and Multi Layer Perceptron.

There were some extra lectures too in the class. They were held usually on Monday or Wednesday and their topics were ones which cannot be dealt with because of time limitation, like Programming in Python and NumPy, Principal Component Analysis (PCA), and Convex Optimization.

Research

I had worked on research with my advisor, David McAllester at TTIC. Since I had only 3 months, I could not achieve something new but learned a lot. The research task was Question Answering where we need to It has relation to Natural Language Processing (NLP), Machine Learning (ML). While I am working on Sentiment Classification at TTI in Japan, I asked Mr. McAllester to give a different task to learn something which is different from Sentiment Classification and appropriate for me to learn in TTIC. I tried implementing a model proposed by Rudolf Kadlec et al. so called Attention Sum Reader. The library I used is TensorFlow. It is the neural network library I am the most familiar with and has a distributed computing feature, and cloud and other commercial support by Google.

TTIC has 2 clusters of Gauda and Slurm. The former one is an older one and an administrator said its resources can be merged into the Slurm cluster. The latter one is a newer one. It is equipped with a bunch of (around 20) computers with multiple CPUs and GPUs. It is a sophisticated environment for research. Anybody in TTIC can use them if their permissions are granted. There are many tools and libraries like OpenMP and Eigen to exploit multiple CPUs in a box. Moreover, some tools and libraries provides convenient ways to deploy programs onto multiple devices and create a distributed system which consists of multiple processes running on them.

I implemented a program as a task of my research work which runs on the Slurm cluster and train the Attention Sum Reader model asynchronously. In asynchronous training of ML models, their parameters are updated asynchronously. Then, we can exploit multiple CPUs and GPUs on multiple machines easily. Another merit of asynchronous training is that models trained asynchronously can be generalized better than models trained synchronously. First, the program prepares datasets for training and evaluation fetching some files over the Internet. Second, it prepares processes of a master and slaves on each node in the cluster to start negotiation before training. The master process coordinates all processes for training and conduct evaluations if requested at the beginning and slaves executes training exclusively. Third, all processes start running their work and cooperate as a big program. I believe that boosts research cycle enormously because we need to train and evaluate a number of models in development of new methods in research. In fact, the program reduced time of a cycle of training and evaluating a model approximately linearly.

How to make use of the experience

Especially on research, I am pretty sure that I can apply knowledge I got at TTIC to my research work. The best knowledge I got there is one about distributed computing and how to exploit multiple GPUs and CPUs over multiple devices in a network or cluster. Although my laboratory does not have any cluster like ones in TTIC, the techniques and tools can be applied. For example, I have developed a command framework built on top of TensorFlow which supports both single and multiple machine, which means that a command of model training or inference is created with it, the command can run on both local machine and distributed system. It is currently beta version but going to be great foundation of my future work.

The experience that I talked with students with top-level knowledge on ML was also great opportunity to enhance my one and improve my conversational skill in English. I made some Chinese friends at TTIC. Honestly speaking, there was some prejudice about Chinese people in my mind. It disappeared after talking with them face to face. We talked about even sensitive topics from political one to historical one, such as Tiananmen Square protests and Nanking massacre. I will be an IT engineer and would like to work abroad in the future. Then, the experience should be helpful.

Even now, it is working effectively for me. I participate some open source projects like TensorFlow daily. They are hosted by some web services like GitHub or BitBucket, where I need communication skill in English to work with other developers all over the world.

In conclusion, the experience that I had learned a lot of Machine Learning, computer programming, English and more must be great help for me in various scenes in my life from now to the future.