

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 on the south side 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 to get off. We had never known such semi-automatic door system. People in Chicago were very kind and helpful.

Our apartment was pretty good and much more than enough honestly. It was much larger than we thought. And, we could see great views of the Michigan lake and Chicago city through windows in our room at (probably) 32nd floor. It also had 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. Its products were a little more expensive than ones in other shops 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 stuff like meat seemed to be even more expensive than ones in Japan. That can be because Chicago is the third biggest city in the US 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 an Asian noodle restaurant called Noodles. I ordered a Singapore noodle there and it had Asian taste and was really good. I also got 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 friends. They sell variety of them, such as hamburgers, hotdogs, 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 and soups.

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. However, 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 work.

In the ML class, we studied about themes of wide range from basic ML to its practical application and some state-of-the-art stuff 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. After that, the themes changed into more concrete ones and their details like logistic regression, SVM, Gaussian mixture model and Neural Network.

The class assignments were exciting. As in Japan, we had some mathematical problems of symbolic calculation and proofs in our homework. In addition to them, we had programming 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 for ML and data science community and holds some public programming competitions of ML and data analytics. We had 4 competitions of multivariate logistic regression of MNIST handwritten digit classification, linear and non-linear SVM of sentiment analysis, Gaussian mixture model, and Neural Network of Multi-Layer Perceptron.

There were some extra lectures in the class too. They were held usually on Monday or Wednesday and their topics were ones which cannot be dealt with in the regular classes 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 of new things, I could never learn in Japan. My research theme there was Question Answering where ML models need to answer a question by reading a document. It has relation to Natural Language Processing (NLP), Machine Learning (ML). While I was working on Sentiment Classification mainly at TTI in Japan, I asked Mr. McAllester to give a different task to learn something different from Sentiment Classification and appropriate for me to learn in TTIC. I had implemented a model proposed by Rudolf Kadlec et al. so called Attention Sum Reader using a Neural Network library, TensorFlow, which can be the most popular one nowadays. It is also a Neural Network library I am the most familiar

with and has some great features like distributed computing function and cloud support.

TTIC has 2 computer 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 in the near future. The latter one is a newer one. It is equipped with a bunch (around 20) of 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 available on it like OpenMP and Eigen to exploit multiple CPUs and GPUs. 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 and preprocessing them. 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 to do so at the beginning, and slaves execute training exclusively. Finally, all processes starts their work and cooperate as a large and distributed system. I believe that boosts research cycle enormously because we usually need to train and evaluate a number of models in the development of new methods in research. In fact, the program reduced time of a cycle of training and evaluating a model approximately linearly to the number of devices I used.

How to make use of the experience

Especially on research, I am pretty sure that I can apply knowledge and techniques I acquired at TTIC to my research work. The best knowledge I got there is one about distributed computing and how to exploit multiple CPUs and GPUs 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 using other common tools. For example, I have developed a command framework built on top of TensorFlow which supports both single and multiple machines after I came back to Japan. Once a command of model training or inference is created with it, the command can run both on a local machine and distributed system. It is currently beta but going to be great foundation of my future research 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 also made some Chinese friends at TTIC. Honestly speaking, there

was some prejudice against Chinese people in my mind. However, it disappeared after talking with them face to face. We talked about even sensitive topics from political one to historical one, such as the Tiananmen Square protests and Nanking massacre. I will be an IT engineer and would like to work abroad in the future. Then, this experience should be helpful to communicate and understand each other with people in other countries.

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