NEC Project:

Title Annealing machine application to artificial neural networks

Industrial Partner NEC Corporation

NEC (https://www.nec.com), founded in 1899, is now particularly addressing the development of solutions for society that will help resolve many issues the world is facing and which will engender the creation of a brighter and more prosperous society. Through co-creation initiatives with many different stakeholders, including customers, business partners, private individuals, government agencies, and international institutions, we are actively devising new business models to create social value by harnessing our extensive information and communication technology (ICT) assets.

Industryial Mentor
Ryoji Miyazaki, Ph.D. NEC Corp.

Background

Machine learning is applied increasingly to various industrial problems including image recognition, natural language processing, and material informatics. A recent breakthrough in this field has been led by deep learning based on many-layered artificial neural networks. One such architecture is a deep belief network [1] that can be interpreted as a composition of Boltzmann machines (BMs). A BM is a network which learns the distribution of data [2]. Actually, BMs are named after the Boltzmann distribution, which BMs assume as a data distribution with latent variables to follow. The distribution was introduced in physics to describe thermal equilibrium. The trained network is expected to generate possible “data’’ distributed similarly to the true data. In training, we adjust weights and biases of the network. We must collect statistics based on a model for comparison with those of the true data. Unfortunately, the sampling process requires huge computational costs. Some algorithms based on a rough approximation have been used to avoid computation time growth. For our project, we attempt the application of another technology, an annealing machine, to this problem of sampling.

Great interest has arisen in developing so-called annealing machines over the last decade because combinatorial optimization, the primary target of the annealing machines, is a difficult, ubiquitous problem among many industries. Annealing machines use some algorithms or heuristics such as simulated annealing (SA), which imitates the gradual decrease of the value of a parameter, called a temperature, to find solutions [3]. Actually, SA moves between states for the problems so that the state distribution converges to the Boltzmann distribution at the temperature. In fact, SA “anneals’’ the model to use the fact that the Boltzmann distribution at low temperatures gives the state which minimizes the cost function for a problem with high probability. A task of generating the Boltzmann distribution lies behind SA. It is noteworthy that we face the same task in training BMs. This fact implies that annealing machines can be used to accelerate the BM training process.

NEC has been developing a system for an SA-based algorithm that works on SX-Aurora TSUBASA [4]. SX-Aurora TSUBASA, the latest model of NEC’s supercomputers, has a vector processor that features extremely high memory bandwidth (1.2 TB/s) and 2.45 TFLOPS performance. NEC has implemented software running a variant of SA that takes advantage of the potential of SX-Aurora TSUBASA. This annealing machine, the system composed of SX-Aurora TSUBASA and the software, works well for its intended task of solving combinatorial optimization problems rapidly. However, we are interested in another application. As described above, annealing machines might help to train BMs. Their application to sampling is a frontier of the study of annealing machines. Although they have been investigated extensively, the study of them has been limited to small BMs. The development of large annealing machines, such as NEC’s supported attempts to apply them for large BMs, has continued, but no report of the literature has described their successful application. If applied successfully, then the application might present quite new alternatives to tackle the notoriously difficult problem of sampling in machine learning.

Project and Expectations

The project is expected to achieve the first application of a large annealing machine, NEC’s system composed of SX-Aurora TSUBASA and software for an annealing algorithm, to the neural network training process, in particular BMs. If their application is successful, then we expect students to demonstrate the benefits and shortcomings of this approach and how it might accelerate standard approaches. Students must develop an algorithm to run BMs using the potential of NEC’s annealing machine. Students will have a unique experience because the system has not been released and because its hardware, SX-Aurora TSUBASA, is a very high performance vector super computer.

Students need not know details of BMs or annealing machines. They can learn topics during this program. Linux experience and Python/C coding skills are necessary. Your skills will be needed to run the annealing machine.

NEC will provide computational environments to students so that they can use the annealing machine on SX Aurora TSUBASA during this program. An example of training datasets will be MNIST [5]: a publicly available database of handwritten digits. Trained BMs can generate artificial images of handwritten digits and can also reconstruct corrupted images. However, we do not restrict students to sole use of MNIST. It would be interesting to use other datasets and to compare the obtained results to ascertain the structures for which the annealing machine shows good/poor learning performance.

Requirements

Linux experience and programming skills (Python and C).

References

[1] G. E. Hinton, S. Osindero, Y.-W. Teh, "A Fast Learning Algorithm for Deep Belief Nets," Neural Computation **18**, 1527 (2006).

[2] G. E. Hinton, "A Practical Guide to Training Restricted Boltzmann Machines," in Neural Networks: Tricks of the Trade, Second ed., Lecture Notes in Computer Science, edited by G. Montavon, G. B. Orr, K.-R. Miller (Springer, New York, 2012), **7700**, 599.

[3] S. Kirkpatrick, C. D. Gelatt Jr., M. P. Vecchi, "Optimization by Simulated Annealing," Science **220**, 671 (1983).

[4] NEC's website on SX-Aurora TSUBASA, https://www.hpc.nec/

[5] Y. LeCun, C. Cortes, C.J. C. Burges (1998) The MNIST database of handwritten digits. http://yann.lecun.com/exdb/mnist/