Overview of Dr Nghiem's research

Introduction

In recent years, with Google, Microsoft Azure and Amazon launching their cloud machine learning platform, machine learning (MI) and artificial intelligence (AI) have been prominently developed. Machine learning is a data analysis method which is built by automated analysis models. It is a branch of artificial intelligence(AI) based on the idea which system can learn from data, identify patterns, and make decisions with minimal human intervention. Machine learning was born from pattern recognition and the theory that computers can learn without programming to perform specific tasks; researchers interested in artificial intelligence want to know if computers can learn from the data.

But surprisingly we have been experiencing machine learning without knowing it. The optimization algorithm is an essential tool for solving parameters. The most important optimization algorithms currently are those that can be used to solve constrained non-linear, non-smooth large-scale optimization problems because these challenging problems are becoming more and more critical in modern machine learning. The optimization algorithms commonly used in machine learning mainly include a genetic algorithm (GA), second-order algorithm, proximity gradient (PG) algorithm, and coordinate reduction (CD) algorithm.

On November 15th, 2018, Dr.Truong Nghiem presented his work to everyone in the INF501 course in Northern Arizona University.Dr.Truong Nghiem's research explores the mathematical models to predict future outcomes, then optimizes a cost function to obtain control inputs on that system. In this essay, I will briefly introduce Dr.Truong Nghiem's research and discuss some critical connections between my research and Dr.Truong Nghiem's research.

Research Topic Overview

Machine learning problems are usually converted into an objective function to solve. Dr. Nghiem's research focuses on learning agents' operations with Gaussian Processes(GP), which is an adaptive mechanism. In general, the inputs which are considered to be a set of random variables mathematically are too difficult to be observed in the Gaussian Processes(GP). Gaussian Processes (GP), as a type of statistical models, are particularly attractive due to their modeling flexibility and their ability to provide probabilistic estimates of prediction uncertainty. However, the optimization problem is typically non-convex and highly demanding and scales poorly with model size. Hence, Professor Nghiem developed a model which name is Data-driven Model Predictive Control (MPC) to predict the optimized gradient information from the historical data based on Gaussian Processes. According to Dr. Truong Nghiem, this approach uses mathematical models to predict future outcomes, then optimizes a cost function to obtain control inputs on that system.[1] GP-based MPC has been developed and applied. This causes unsatisfactory solving performance, even with state-of-the-art solvers, and makes the approach less suitable for real-time control. Research team develop a method based on a new concept, called linearized Gaussian Process, and sequential convex Programming, that can significantly improve the solving performance of GP-based MPC. MPC-GP models have constrained versions of the general GP. By imposing model structure (known fixed aspects and know random aspects), known parameter interactions, and parameter limits, a least square estimate for each time step can be produced. For example, control systems which are used to monitor and supply heating, cooling, lights, and electricity can be built efficiently and effectively with MPC. Control systems depend on a model's prediction of outcomes based on various environmental factors, the requirements of the end user, and the state of the system being controlled. Based on the predictive variance, the coordinator decides whether to communicate with the agent or not.

Discussion

Although my research topic which is pair programming based on computer programming self-efficacy is not very close to machine learning (MI) and artificial intelligence (AI), I also can distinguish some critical connections between my research and Dr.Gerosa’s research.

Both of us used an optimization algorithm to solve the problems in each of our studies. The genetic algorithm (GA) is an optimization method to solve constrained and unconstrained optimization problems based on natural selection. In my research, I used a genetic algorithm (GA) to find the minimum difference of the self-efficacy score between two students in a pair and also want to minimize the times of two students working together. By repeatedly modifying a population of individual solutions in each step, the genetic algorithm randomly selects solutions from the current population to be parents and use them to produce the children for the next generation. In Dr. Nghiem’s research, the research team developed a model to predict the optimized gradient information from the historical data based on Gaussian Processes. This predictive model was developed based on the linearized Gaussian Process to deal with nonconvex optimization problems which are rather typical. The efficiency and advantages of both predictive models are demonstrated using critical numerical examples.

The genetic algorithm(GA) in my research doesn’t have an inherent notion of best match progress. If our program is running for a long time, we expect our best-stored solution to be better than after a short run, but we can’t quantify how good our current answer is, relative to abound on the global optimum. If the genetic algorithm(GA) doesn't encounter a lucky mutation(solution), the genetic algorithm doesn't have a way to conduct a perfect solution. In contrast, the Gaussian Processes(GP) doesn't seem to have this problem. The Gaussian Processes(GP) seem to address this issue in a statistical sense - the variance of the acquisition function at a point in the search space gives an idea of how thoroughly that region has been searched, allowing the search to guide itself out of a large shallow bowl, given enough runtime.

A further benefit of Dr. Nghiem's GP approach to MPC is that the model is more scalable and predictable, which makes it more suitable for real-time control. Because this approach is based on the concept of linearized Gaussian Process (linGP), proposed in this paper, and Sequential Convex Programming. Hence, even the datasets grow more substantial; the researcher also can easily predict the results. And this approach not only solves GP-MPC faster than other NLP methods but also is much less influenced by the GP training data size – a key factor affecting the computational complexity of GPs and GP-MPC.

Conclusion

MPC-GP and linGP-MPC can be extremely useful. For some complex problems and large datasets problems with simple objective functions, evaluating the quality of an individual using the objective function is prohibitively expensive. By building a statistical model based on earlier observations, the GP-MPC approach can reduce the number of redundant evaluations of the objective function, making it more efficient.

Reference

[1] W. Shi, Q. Ling, G. Wu, and W. Yin. Extra: An exact first-order algorithm for decentralized consensus optimization

[2] A. Nedich, A. Olshevsky, and W. Shi. Achieving geometric convergence for distributed optimization over time-varying graphs

[3] J. Chen and A. H. Sayed. Diffusion adaptation strategies for distributed optimization and learning over networks