Predicting Latency in 5G Networks by Using Machine Learning

Erica Elgcrona & Evrim Mete

June 13, 2023

Our work has shown that it is possible to predict latency in 5G networks 4 seconds in advance with good accuracy. This was done using a machine learning approach where four different models were used. The best performing model was the CNN model and all models were improved using Bayesian optimization.

Imagine being at home watching a live soccer game and the stream freezes just seconds before the final goal. We all know how frustrating that can be. Should it still be a problem in 2023? Thanks to the latest technological developments in telecommunications and networking, it may not be an issue anymore. Companies that provide streaming services could prevent it by working on latency prediction in their network. When streaming video, information is sent across a network in small "packets." As these packets move across the network, there can be a delay, which is known as latency [1].

The latest 5th generation network, 5G, offers even faster and more reliable transmission than previous generations [2]. Even if you are not physically present at the stadium, the disparity between watching the game from the comfort of your home via a streaming service will be reduced. Although the 5G protocol itself has already been designed with these new enhancements, specific adaptations could be made for specific user scenarios, such as predicting latency in advance to minimize its impact. There are several possible methods for predicting latency, and one of the most exciting is machine learning. Machine learning uses programming techniques inspired by the human brain to teach machines to perform specific tasks without explicit coding.

Our work has shown promising results on the possibility of using machine learning to detect the magnitude and timing of latency peaks, by analyzing a large data set from Sony's 5G lab. The project used four different popular model architectures. These models were Forward Neural Network (FNN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Long Short Term Memory (LSTM). All of these models contain different layers with different mathematical operations, and the different types of layers approach the problem from different perspectives. Each layer also contains different hyperparameters, which can be thought of as specific settings, that can be adjusted to model the data as good as possible.

After testing the initial models, an approach called Hyperparameter Optimization (HPO) was used. This is a standardized method for optimizing the hyperparameters in different models [3]. There are several different methods for HPO that can be more or less efficient. Specifically, the method we used for HPO was Bayesian optimization. This approach starts by randomly choosing a set of parameters and evaluating their performance. Based on how good these results are, the algorithm decides on a new set of parameters and repeats this process until it finds the best possible set of parameters [4]. By using Bayesian optimization, we were able to improve our models even further, as shown in Figure 1.

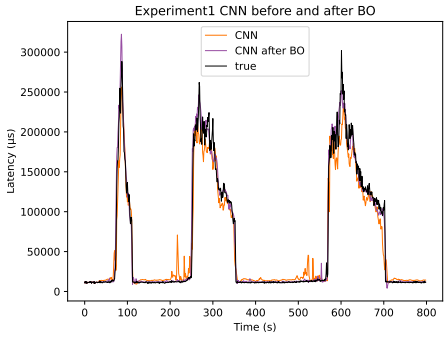

Figure 1: Results for the CNN model before and after Bayesian optimization, along with the true latency. The model performs better after Bayesian optimization.

References

[1] Kurose, J. F., and Ross, K. W. (2017). Computer Networking: A Top-Down Approach (7th edition). Pearson.

[2] Allen, J.. (2020). What about 5g? Road Warrior/ GP Solo 37:6, pp. 85–86. ISSN: 1528 - 638X

[3] Snoek, Jasper and Larochelle, Hugo and Adams, Ryan P. (2012). Advances in Neural Information Processing Systems. Practical Bayesian Optimization of Machine Learning Algorithms. Curran Associates, Inc.

[4] Aghaabbasi, Mahdi and Ali, Mujahid and Jasiński, Michał and Leonowicz, Zbigniew and Novák, Tomáš. (2023). On Hyperparameter Optimization of Machine Learning Methods Using a Bayesian Optimization Algorithm to Predict Work Travel Mode Choice.