ML in Telecommunications: Building Intelligent Systems with Machine Learning

**Name: Christ-Brian Amedjonekou**

**Date: 05/22/2019**

**TCET 3102-E316 (Analog and Digital Communications) Final Report: ML in Telecommunications**

**Spring 2019, Section: E316, Code: 37251**

**Instructor: Song Tang**

# **Abstract**

In this report, we set out to explain how Machine Learning is used to improve Telecommunication systems, particularly with mobile networks. The goal of this report is to give a brief history/background of Machine Learning, explain how it’s used in the Telecom industry, and describe the characteristics/features of Machine Learning. The report will also provide comparisons between Machine Learning in Telecom and other similar methods. This report will also go over the uniqueness of Machine Learning as well as the possible improvements/growth it can undergo in Telecom. Machine learning is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on patterns and inference instead. Overall, we will demonstrate how effective machine learning techniques are for performance prediction problems in wireless networks.

**Introduction**

What is Machine Learning? One technical definition includes the following: Machine learning is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on patterns and inference instead. The goal of Machine Learning is to be predictive (some, usually business folk, refer to machine learning as being predictive analytics). Machine Learning accomplishes this by using several borrowed Mathematical Modeling techniques including, but not limited to, Mathematical Optimization which is the selection of a best element (with regard to some criterion) from some set of available alternatives. The main difference in approach between more traditional Mathematical Modeling paradigms and Machine Learning is the fact that Machine Learning very Data Driven as opposed to being Analytical Driven. This difference, which we will discuss in detail in later sections, gives us an advantage in modeling/predicting real world implications. For now, we’ll discuss how the use of Machine Learning all began.

**History/Background**

While Machine Learning may seem relatively new, it is simply the combination of math known for centuries plus the advancements in computing over the past 70 years that form the foundation of what we know Machine Learning to be today. Many of the mathematical underpinnings of modern machine learning predate computers and come from statistics. We’ll separate the history/background of machine learning by decade starting from the 1940s. But before we begin, let us visit pre-1940s history. Bayes Theorem, a theorem developed by Thomas Bayes in the 18th century about conditional probability, is foundational to Machine Learning; A Machine Learning model call Naïve Bayes Classifier trains Supervised Models using Bayes Theorem. Regression, another important method used in Machine Learning, uses the Least Square Method for data fitting which was developed by Adrien-Marie Legendre. Finally, Markov Chains, developed by Andrey Markov and used in Stochastic (Probabilistic) Modeling, is also a major contributing factor towards the development of Machine Learning.

The 1940s, particularly 1948, was a very important timeframe. It can be said, and it would not have been overstated, that what occurred in the 1940s would set off the modern computer revolution. In the 1940s, particularly 1948, stored-program computers that could hold their instructions (programs) in the same memory used for data had made significant progress. In 1948, the Manchester Small-Scale Experimental Machine, also nicknamed 'Baby', was developed. Similar machines were also developed in 1949 (Cambridge’s EDSAC and the Manchester Mark 1) and 1951 (University of Pennsylvania’s EDVAC) respectively. Being able to store information of any kind is an essential task in Machine Learning.

In 1950, Alan Turing, considered by most to be the godfather of theoretical computer science, published *Computer Machinery and Intelligence* where he posed the following question: “Can machines think?”. It would become one of the first papers to attempt to describe how ‘artificial’ intelligence could be developed; Machine Learning is a subset of Artificial Intelligence. In the paper, Turing proposed an imitation game. The goal of the game was to determine whether a computer was intelligent. This was done by asking a person to distinguish between a human and a computer when communicating with them both through typed messages. The idea was that if the individual asked to distinguish between two could not, then the machine could be considered intelligent. Alan Turing published paper, *Computer Machinery and Intelligence,* was not the only significant advancement made towards the development of Machine Learning (Artificial Intelligence as a whole) during the 1950s. A year later, in 1951, Marvin Minsky and Dean Edmonds built the first artificial neural network. An Artificial Neural Network, abbreviated ANN, is our take on another powerful computing device, our brains. An ANN, shown in Figure 1 below, models a network of synapses in the brain.

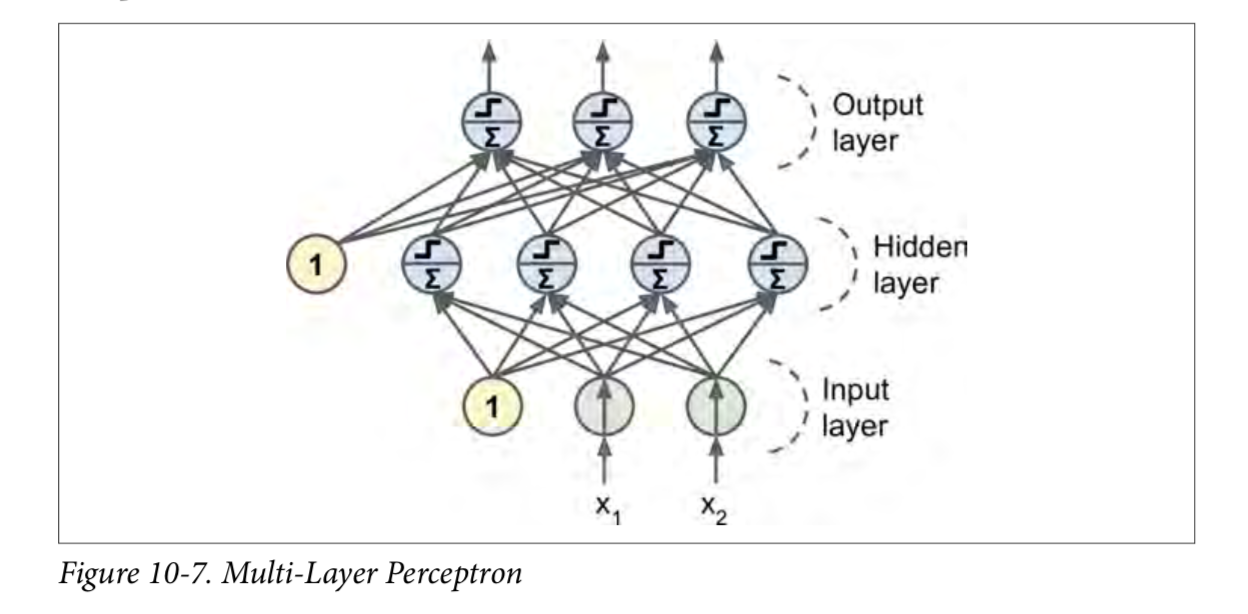


Figure 1: Multi-Layer Perceptron (Simple ANN)

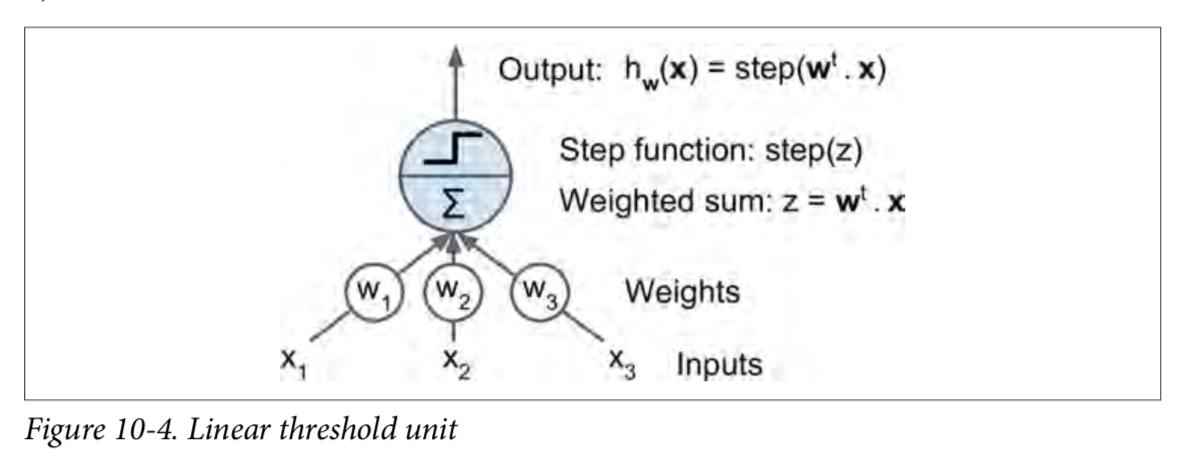


Figure 2: Linear Threshold Unit (Model of a Perceptron)

Each node in the graph (network), shown in Figure 1, is a rough model of the operation of synapses in the brain called Linear Threshold Units (Figure 2); Linear Threshold Units (LTU) is another way to refer to the Perceptron. Each Perceptron is responsible computations and connecting them together allows you to model, theoretically, every possible operation imaginable. The most important feature was that the artificial neural network could learn from experience. These findings led to unrealistic expectations when it came to Machine Learning. As a result, during the 1960s and 1970s Machine Learning (and Artificial Intelligence) received a major decrease in funding. This event became known as the First A.I. Winter.

In the 1980s and 1900s, Machine Learning was on the receiving end of a resurgence. That was because by the 80s and 90s computational power had significantly improved several folds. Additionally, new approaches in machine learning, the emergence of expert systems, and the rediscovery of old ideas that could be applied in new settings all contributed to the resurgence. One specific event, namely the IBM computer Deep Blue beating world chess champion Garry Kasparov, once again pushed Machine Learning to the forefront. Deep Blue was able to accomplish this in part because it relied on brute computing power. Deep Blue worked by searching 6 to 20 moves ahead at each position, having learned by evaluating thousands of old chess games to determine the path to checkmate.

2006 was another important date. An essential Machine Learning Technique used to train Deep Neural Networks, Artificial Neural Networks with more than one hidden layer, called backpropagation was developed. Backpropagation works by applying the following:

* Each instance of the training set is fed to the neural network.
* The neural network then computes the output of every neuron in each consecutive layer (this is the forward pass, just like when making predictions).
* It then measures the network’s output error (i.e., the difference between the desired output and the actual output of the network), and it computes how much each neuron in the last hidden layer contributed to each output neuron’s error.
* Then, it proceeds to measure how much of these error contributions came from each neuron in the previous hidden layer—and so on until the algorithm reaches the input layer.

Basically, it trains the neural network on data and then goes back (in reverse) and checks how well each neuron did its calculation.

As of today, Deep Neural Networks are a staple of the Machine Learning community. They are used by Amazon, Google, YouTube, etc. In 2014, DeepMind Technologies (a Google Subsidiary) gained prominence when it developed a neural network that could learn to play video games simply by analyzing the behavior of pixels on a screen. DeepMind was also responsible for creating the Neural Turing Machine: neural network that can access external memory. In 2016, AlphaGo (also developed by DeepMind), considered to be Deep Blue’s successor, beat the world’s number two Go player Lee Sedol. It followed up that win with another against the number one player Ke Jie in 2017. It uses a Monte Carlo tree search algorithm to find moves. Accomplishments such as these shows how far Machine Learning has come. They also give rise to thought of how far Machine Learning can go. In the next section, we’ll discuss applications of Machine Learning in Telecom.

# **Applications of Machine Learning in Telecom**

Machine Learning in Telecom can be broken down into two categories: application of machine learning techniques for performance prediction problems in wireless networks, and automation of next generation wireless cellular networks through the application of Deep Neural Networks. We’ll being with using Machine Learning to predict performance of mobile networks. The reason why Telecoms are interested in using Machine Learning in the businesses is because of the possibility of automating the optimization and management of wireless mobile networks. The reason to do this is to reduce the operational costs for network operators, as well as to improve the quality of user experience. When we talk about network management today, most of it involves human interaction ranging from conducting drive tests in order to evaluate the network coverage and performance to diagnosing customer complaints. Machine learning techniques can reduce the stress that accompanies these tasks by reducing the need for drive tests and helping to predict and diagnose network failures even before they noticeably degrade the quality of service of the network users. As aforementioned, Machine Learning is primarily data driven. We’ll briefly go over an example problem and discuss how Machine Learning can make an impact (*and be predictive*).

In the example, the data set (shown in Figure 3), corresponds to a 10 km × 10 km subset within a large professional drive test campaign conducted in a major metropolitan area in the US. Although the full data set covers an area of approximately 120 km × 170 km and consisted of 4 million measurement locations in total, the good thing about Machine Learning is that it can learn. Therefore, it can be scaled. The gathered measurements are all for the performance characteristics of a 2 GHz downlink band for a major 4G cellular network operator. The variables included in the data set are geographical coordinates and time of the measurement, the received signal strength (RSS) measured at each location in dBm, the carrier to interference (C/I) ratio measured in dB, as well as the achieved user data rate (in bps) for a simple data download application. Information on the velocity of the mobile drive test unit is also added, as it is expected to play a role especially in the achieved data rate.

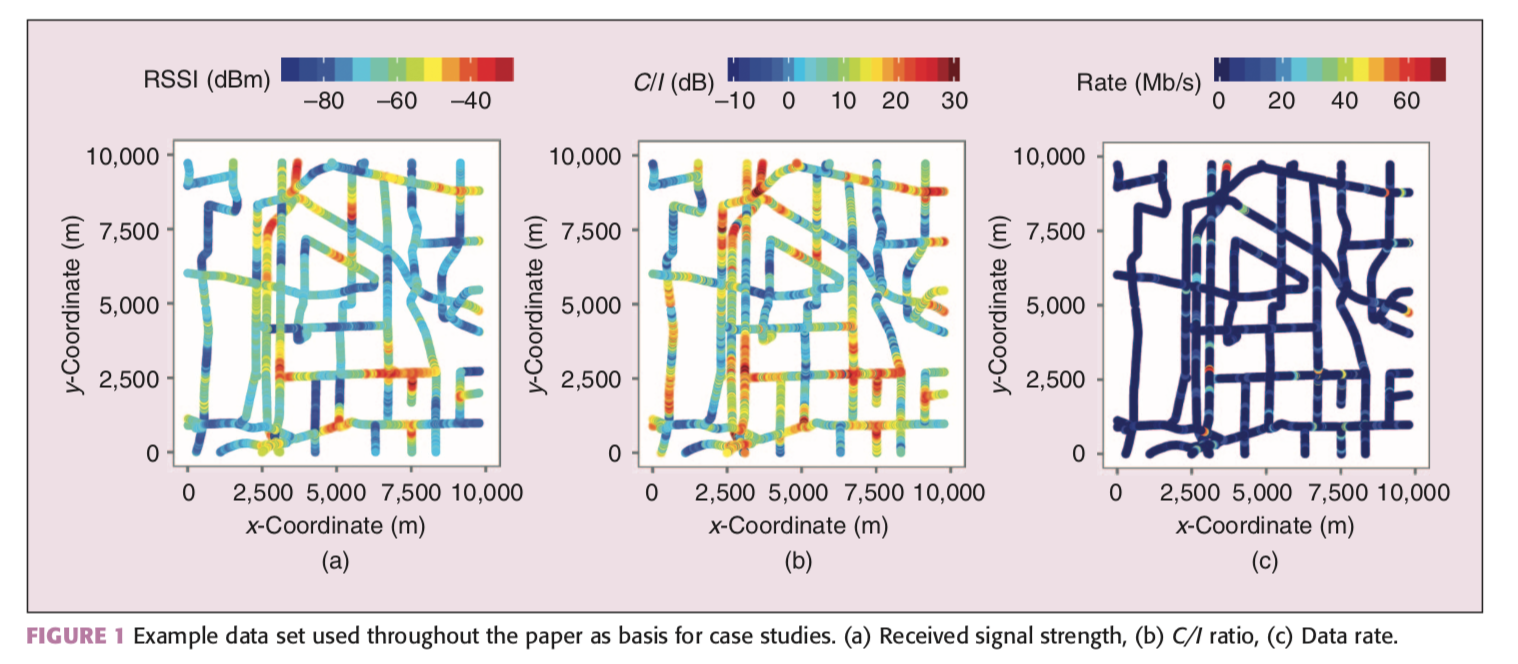


Figure 3: Example Dataset Plotted

One of the potential problems where Machine Learning can be useful is in approximating well the Shannon capacity formula which tells the maximum rate at which information can be transmitted over a communications channel of a specified bandwidth in the presence of noise. Regression, particularly Linear Regression, can be used to solve this type of problem. The advantage of using Linear Regression is that it is simple to use. Figure 4 shows the math behind Linear Regression. The process in which the model is implemented is as follows:

* First you would examine your data and determine if cleaning is necessary to get it in the proper form to implement the model
* Then you would visualize your data in some form of plot (as shown above) to see if you can determine what model would fit best to the problem.
* Then you would split your data into two sets: a test set, and a training set. The training set is used to train the model as the name suggest. The test set is used to see if the model selected performs well on new data.



Figure 4: Linear Regression

The coefficients are determined by minimizing the root mean squared error (RMSE) of the predictor. ‘y’ becomes simply the best approximation of the network throughput as the optimal weighted sum of the individual Shannon capacity estimates. As a result, linear regression can be used as an improved proxy for simple SINR-based user data rate models relying on Shannon capacity formula.

Another potential problem where Machine Learning can be useful is in addressing the issue of HO prediction in mobile networks. A handover is a process in telecommunications and mobile communications in which a connected cellular call or a data session is transferred from one cell site (base station) to another without disconnecting the session. Markov Chains can help in the following ways:

* Modeling Mobility behavior of the users with Renewal Markov Renewal Process (MRP) to predict both the single and multi-transition.
* hybrid Markov-based models can be used for human mobility predictions

An example of a Markov Chain is shown in Figure 5.

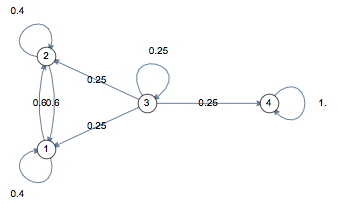


Figure 5: Markov Chain

Artificial Neural Networks can be used to learn the user patterns to use them for future location predictions in UMTS networks, decreasing the location update signaling overhead.

# **Conclusion**

Machine Learning is an emerging field as of 2019. In a society where we have massive amounts of collected data, it can be quite effective to use a data driven approach. This is especially true in the Telecom industry. Machine Learning can be quite effective with its predictive qualities. As such it is no surprise that its use is being adopted by Telecom in order to optimize our mobile networks. We can’t wait to see how for it can go.

# **Citations**

1. J. Riihijarvi and P. Mahonen, "Machine Learning for Performance Prediction in Mobile Cellular Networks," in IEEE Computational Intelligence Magazine, vol. 13, no. 1, pp. 51-60, Feb. 2018. doi: 10.1109/MCI.2017.2773824 keywords: {cellular radio;Gaussian processes;learning (artificial intelligence);regression analysis;mobile cellular networks;performance prediction problems;network performance;direct measurements;real-world drive test data;Gaussian process regression;wireless mobile networks;Machine learning;Predictive models;Cellular networks;Wireless networks;Mobile computing;Machine learning algorithms;Learning systems;Guassian processes}, URL: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8253758&isnumber=8253704>
2. Wikipedia contributors. (2019, April 1). Shannon–Hartley theorem. In Wikipedia, The Free Encyclopedia. Retrieved 22:23, May 23, 2019, from <https://en.wikipedia.org/w/index.php?title=Shannon%E2%80%93Hartley_theorem&oldid=890534229>
3. Wikipedia contributors. (2019, May 23). Mathematical optimization. In Wikipedia, The Free Encyclopedia. Retrieved 22:23, May 23, 2019, from <https://en.wikipedia.org/w/index.php?title=Mathematical_optimization&oldid=898431603>
4. Wikipedia contributors. (2019, May 23). Machine learning. In Wikipedia, The Free Encyclopedia. Retrieved 22:24, May 23, 2019, from <https://en.wikipedia.org/w/index.php?title=Machine_learning&oldid=898379582>
5. Mann, G. (2018, May 03). Machine Learning vs. Mathematical Modelling in Practice. Retrieved from <https://www.leansystems.co/blog/machine-learning-vs-mathematical-modelling>
6. Mann, G. (2018, May 03). Machine Learning vs. Mathematical Modelling in Practice. Retrieved from https://www.leansystems.co/blog/machine-learning-vs-mathematical-modelling
7. A novel deep learning driven low-cost mobility prediction approach for 5G cellular networks: The case of the Control/Data Separation Architecture (CDSA). (2019, January 17). Retrieved from https://www.sciencedirect.com/science/article/pii/S0925231219300438