Research Automated Machine Learning

Shaun Pritchard

Rasmussen College

QMB3000

Kevin Watts

May 2, 2020

**Research Automated Machine Learning**

**Definition of machine learning:**

Machine learning is actually a branch subset of artificial intelligence. Artificial intelligence (AI) is a wide-ranging branch of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence. There many different subsets of artificial intelligence. Machine learning also has different subsets such as deep learning. Using Computing we can get to design systems that can learn from data in a matter of being trained.

Machine learning is the use of mathematical procedures such as algorithms and statistical inferences to analyze data. With the aim to discover useful patterns, relationships, correlations, and trends. All of which provide real information, solutions, and applications in the real world. In 1959, Arthur Samuel defined machine learning as, “A Field of study that gives computers the ability to learn without being explicitly programmed.” Samuel is credited with creating one of the self-learning computer programs with his work at IBM. *(Carnegie Mellon, 2020)*

There are a number of different algorithms that we can employ with machine learning. The required output of our data inferences is typically what we use to decide which machine learning algorithms to use. We can select which algorithms to choose based on the characteristics of two learning types in which we use to train data sets. These learning types are known as supervised or unsupervised. When we get into the world of deep learning then we can have semi super where we have a mix of both.

Supervised learning refers to working with a data set of label training data forever example in our data we typically have an input object an output object. For instance, if we wanted to use machine learning to filter emails that were spam from good emails. We would have to label a set of emails that were considered spam and use an algorithm that could look for the specific traits of those emails to classify new incoming emails and determine if they need to go to the spam folder.

Whereas with unsupervised learning is the complete opposite we let our algorithms find hidden patterns in the data. The algorithms would then create training models that we can use to test the algorithms until we get the desired output or inferences. Then we can take the trained models and put them into production to handle the data sets input on real data. With unsupervised learning, there is no wrong or right answers just a case of running machine learning algorithms and seeing what patterns and outcomes occur.

Machine learning as essentially changes the world we live in and we see it every day most people unknowingly do not even realize it. For an example of Google search engine algorithms. They process hundreds of data points off of each user's profile set to display specific results based off of past history recent searches products purchased many other factors to determine what information to show the searcher. Machine learning algorithms are responsible for compiling all these data points in determining the output decisions *(Miller, 2019)*.

There are many other features and capabilities for machine learning as follows:

* Voice recognition
* Computer vision
* software
* Spam detection
* Search engines
* Stock trading
* Robotics
* Medicine and health care
* Advertising
* E-commerce
* Retail
* Game Analytics
* Weather patterns
* Political trends

The list goes on and on, and we can use a number of the different programming languages to implement machine learning algorithms and applications such as:

* Python
* R
* Matlab
* Octave
* Scala
* Java
* And many web-based languages such as JavaScript and Ruby

We can use these tools to plan projects and run inferences by implementing a specific cycle of actions that need to be performed. First, we acquire and correlate the data. Then we prepare the data by cleaning it checking it for quality. Followed by the process I am running Machine learning Tools and programs. finally reporting the data in results and presentations for visualization or information for making business type decisions. Just as a side note I like to add that we can test inferences and output visuals in 3D space with digital objects with gaming engines such as unity. As well as implement an array of visualization products such as Power Bi, Azure, Tabluee, and excel *(Spotless Data, 2020)*.

**General challenges of applying machine learning**

With machine learning we do have many advantages in which we benefit. With these advantages comes some very prominent challenges as well.

**Advantages of Machine learning**

***1. Easily identifies trends and patterns***

Machine Learning can review large volumes of data and discover specific trends and patterns that would not be apparent to humans. For instance, for an e-commerce website like Amazon, it serves to understand the browsing behaviors and purchase histories of its users to help cater to the right products, deals, and reminders relevant to them. It uses the results to reveal relevant advertisements to them.

***2. No human intervention needed (automation)***

With Machine Learning, you do not need to babysit your project every step of the way. Since it means giving machines the ability to learn, it lets them make predictions and also improve the algorithms on their own. A common example of this is anti-virus software; they learn to filter new threats as they are recognized. Machine Learning is also good at recognizing spam.

***3. Continuous Improvement***

As Machine Learning algorithms gain experience, they keep improving in accuracy and efficiency. This lets them make better decisions. Say you need to make a weather forecast model. As the amount of data, you have keeps growing, your algorithms learn to make more accurate predictions faster.

***4. Handling multi-dimensional and multi-variety data***

Machine Learning algorithms are good at handling data that are multi-dimensional and multi-variety, and they can do this in dynamic or uncertain environments.

***5. Wide Applications***

Machine Learning holds the capability to help deliver a much more personal experience to customers while also targeting the right customers.

**Disadvantages of Machine Learning**

With all those advantages to its powerfulness and popularity, Machine Learning is not perfect. The following factors serve to limit it:

***1. Data Acquisition***

Machine Learning requires massive data sets to train on, and these should be inclusive/unbiased, and of good quality. There can also be times where they must wait for new data to be generated.

***2. Time and Resources***

Machine Learning needs enough time to let the algorithms learn and develop enough to fulfill their purpose with a considerable amount of accuracy and relevancy. It also needs massive resources to function. This can mean additional requirements of computer power for you.

***3. Interpretation of Results***

Another major challenge is the ability to accurately interpret results generated by the algorithms. You must also carefully choose the algorithms for your purpose. This also means bigger learning curves.

***4. High error-susceptibility***

Machine Learning is autonomous but highly susceptible to errors. Suppose you train an algorithm with data sets small enough to not be inclusive. You end up with biased predictions coming from a biased training set. This leads to irrelevant advertisements being displayed to customers. In the case of ML, such blunders can set off a chain of errors that can go undetected for long periods of time. And when they do get noticed, it takes quite some time to recognize the source of the issue, and even longer to correct it.

**Problems that machine learning can solve**

Machine learning is a technological advancement of our human species. Its applications have been implemented and used in a wide variety of tools and achievements. here are a few applications in which currently use machine learning.

Using machine learning, robots can acquire skills or learn to adapt to the environment in which they are working. Robots can acquire skills such as object placement, grasping objects, and locomotion skills through either automated learning or learning via human intervention with self-adaption. Robotics have been adopted by military and police forces to save lives and spare the cost of human lives *(Jamshidi, 2019)*.

About 70 percent of trades are performed by a machine and not by humans on the trading floor. This is all very well when things are going fine, but when a problem occurs it can be minutes before the fault is noticed, by which time many trades have happened. The flash crash in May 2010, when the Dow Jones industrial average dove 600 points, is a good example of when this problem occurred *(Prathipa, 2020)*.

Apple’s Siri service that is on many iOS devices is another example of software machine learning. You ask Siri a question, and it works out what you want to do. The result might be sending a tweet or a text message, or it could be setting a calendar appointment. If Siri can’t work out what you’re asking of it, it performs a Google search on the phrase you said. Siri is an impressive service that uses a device and cloud-based statistical model to analyze your phrase and the order of the words in it to come up with a resulting action for the device to perform. Also, Google has its own android & cloud-based voice search "okay Google" which performs the same actions as SIRI. *(Nguyen, 2019)*

Machine learning is powerful technology and driving force of technology and the world we live in. The benefits of machine learning are exponential and have the ability to advance our civilizations as never seen before.

# References

Brisaboaa, N. R. (2015). Rank-based strategies for cleaning inconsistent spatial databases. *International Journal of Geographical Information Science*, 26.

Carnegie Mellon. (2020). *Arthur Samuel's definition of Machine Learning*. Retrieved from https://www.contrib.andrew.cmu.edu/: https://www.contrib.andrew.cmu.edu/

Cragg, J. g. (1994). Making good inferences from bad data. *Canadian Journal of Economics*, 26.

Greene, K. B. (2015). Bad Data States are embracing the promise. *Governing*, 9.

Grime, L. J. (2013). How Much Does Bad Data Cost Your Company? *Journal of Corporate Accounting & Finance* , 4.

Guess, A. (2011). Root Causes of Data Quality Problems. *Journal of Information Science and Engineering* , 1.

Jamshidi, P. (2019). Machine Learning Meets Quantitative Planning:. *2019 IEEE/ACM 14th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS)*, 13.

Miller, J. D. (2019). *Hands-On Machine Learning with IBM Watson : Leverage IBM Watson to Implement Machine Learning Techniques and Algorithms Using Python.* ebook: Packt Publishing.

Nguyen, T. L. (2019). AI Deep Learning with Convolutional Neural Networks. *Journal of Strategic Innovation & Sustainability*, 12.

Prathipa, J. A. (2020). Stock Market Analytics: Statistical and Machine Learning Techniques. *International Journal of Psychosocial Rehabilitation*, 5.

Spotless Data. (2020). *14 CAUSES OF DATA QUALITY PROBLEMS*. Retrieved from Spotlessdata.com: https://spotlessdata.com/blog/14-causes-data-quality-problems-part-1

Syniti. (2014, 3 24). *The Three Root Causes of Poor Big Data Quality*. Retrieved from syniti.com: https://resources.syniti.com/syniti-blog/the-three-root-causes-of-poor-big-data-quality

W., L. Y. (2006). Journey to Data Quality. *American Psychological Assoc*, 1.