Jagroop Singh

12/24/24

Machine Learning: An Automated Future

Geofrey Hinton, the ‘Godfather of AI’, once stated, “In the long run, curiosity-driven research just works better. Real breakthroughs come from people focusing on what they're excited about.” Artificial intelligence is a field that is rapidly growing and as society approaches the quarter mark of the twenty-first century, one must ask, ‘Will the remainder of this century be dominated by the expansion of artificial intelligence?’ Within the field of artificial intelligence, there exist many subfields of study including genetic algorithms and natural language processing. However, the most consequential field of study is machine learning due to its raw power and incredible scope. Although there are many differing opinions on who coined the term ‘machine learning’, it is undeniable that this field of study began to gain traction during the 1950s and the yearning to understand how a machine can become a better version of itself has only increased since. Machine Learning is an incredible advancement made by society, however, the power that this technology holds can only be harnessed for good if its weaknesses are addressed and if responsible individuals such as developers, engineers, and stakeholders place the necessary safeguards on the development, deployment, and ethical use of this technology.

Without a doubt, research shows that machine learning is a technology that holds immense potential to address serious issues such as making healthcare more efficient, keeping communities safe, optimizing various sectors of the transportation industry, and so much more! Predictive modeling is an extremely helpful process of building models based on available data that can be used to predict potential occurrences, prevent certain outcomes, and quantify the effects that certain measurable variables have on society. Regression testing is a sub-feature of

predictive modeling that can be used to test the credibility of models based on continuous target variables and the principle of finding the line of best fit. According to *Harvard Business Review,* regression testing is important for determining which variables in a situation are crucial for understanding a correlation and which variables hold less significance. For example, in the case of addressing homelessness, regression testing would be a very helpful tool for social workers to utilize in their effort to make sure that the models that they create to solve the issue place heavy emphasis on the more important variables compared to the less impactful variables. Unsupervised learning is an aspect of machine learning in which data is processed with no labels to find patterns within the data without any persuasion from a human. A popular form of unsupervised machine learning is the k-means clustering algorithm which aims to group data points based on their distance from a select few central data points, called 'centroids'. According to Vrata Tech Solutions, the k-means algorithm can play a crucial role in detecting outlier data points which can be translated into scanning for any health anomalies or even as a security tool to keep communities safe. All in all, the positive effects that machine learning can have on society is wide-ranging and its optimization will only lead to more positive results in the future.

Just like any other emerging technology, machine learning has shortcomings that prevent it from being considered a 'perfect tool' to fix every issue faced by society. To begin with, a large portion of successful machine learning models depend on large, specified amounts of data which can lead to unintentional reinforcement of biases held by the data collectors. According to IBM, "Severe or repeated instances of biased or inaccurate AI-driven decisions might spur individuals and communities to question the integrity of the organization deploying the AI." This point singled out by IBM explains the repercussions of an organization deploying untrustworthy models as they can lead to lessened trust by the public and can diminish the fervor surrounding

artificial intelligence. Furthermore, machine learning models often depend on high-quality, rich datasets which require a lot of manual data cleaning, attention to minut details, and access to computer training. The resources and time required to ensure that a machine learning model is prepared to conduct its assignment is very significant and not something that can be overlooked, especially by large organizations who input wide-reaching datasets. While many work to harness the immense power held by machine learning models, it is equally important to recognize this technology's shortcomings to explain any potential blindspots and mitigate their impact. By addressing these limitations in a conscious manner, organizations who want to maintain public trust and brand value can ensure that machine learning models are both effective and ethically credible in their processes.

Next, within the field of machine learning, it is important to make a clear distinction between general machine learning and automated machine learning. While traditional machine learning requires human involvement in the data cleaning and model-choosing processes, automated machine learning automates most traditional steps in an effort to allow engineers to focus more on the analysis of the model. According to *Medium,* while automated machine learning is more efficient and able to define the best model, its largest drawback lies in the fact that engineers have very little control of the parameters used in each automated model. Therefore, automated machine learning should be used when time is a limiting factor and where the problem being measured is simple and its variables clearly defined. Therefore, automated machine learning can be considered an aspect of machine learning that promotes quick models for easily understandable problems. This approach allows organizations to quickly generate insights without requiring large-scale teams with expertise in machine learning techniques.

In conclusion, machine learning is a constantly emerging technology that takes on many forms and can be helpful in a diverse range of fields. In order to truly optimize the positive effects that machine learning models can have on society, it is important to understand the dependencies that this subfield of artificial intelligence has on human involvement through data selection, model picking, and variable tuning. Within this technology exists an automated learning model which limits human intervention, but still contains its drawbacks. Machine learning is a technology that holds tremendous importance in the development of society as it has the power to not only highlight issues but define solutions based on data and facts.