RESEARCHERID
THOMSON REUTERS

[Kumar, 5(8): August 2018]   ISSN 2348 – 8034
Impact Factor- 5.070

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

THE EVOLUTION OF MACHINE LEARNING: APPLICATIONS AND CHALLENGES

Shubham Kumar
Mahatma Gandhi Central University
Email- shubhamkumar@mgcub.ac.in

ABSTRACT

This paper explores the development of Machine Learning (ML), from its historical background and current usage in applications to some of the challenges hindering its more mainstream adoption. It has spurred a change in the industry by allowing machines to learn and improve their performance in the absence of explicit programming. This paper will track the evolution of ML, starting from its origins in the 1950s to the present day, focusing on deep learning and neural networks. Important applications in healthcare, finance, transportation, and entertainment will be discussed and analyzed to illustrate how ML has revolutionized diagnostics, fraud detection, and autonomous driving. However, it is also haunted by a series of challenges in data quality, computational demand, algorithmic bias, and lack of interpretability. It highlights these barriers and looks at potential solutions that could be employed to overcome the named barriers. The findings reveal the importance of addressing such issues in unleashing the true potential of ML, which is largely becoming ethical and responsible in various domains.

Keywords*: Machine Learning, Deep Learning, Applications, Computational Demands, Algorithmic Bias, Interpretability*

I. INTRODUCTION

BACKGROUND TO THE RESEARCH

Machine Learning (ML), the predominant area of AI, is about the development of algorithms that enable systems to improve performance through learning from data without explicit programming. The idea started in the 1950s, with Alan Turing, who first presented the idea of machines mimicking human thinking. The term "machine learning" was coined by Arthur Samuel in 1959 when he developed a program that enabled a computer to improve its checkers-playing abilities based on past games (Rumelhart et al., 1986).

Early computer-based learning was restricted to limited data and power of computing. The improvement of power of computing, large datasets, and algorithmic innovations has led to gigantic breakthroughs, especially deep learning (Zhou et al., 2017). Deep learning employs multiple layer neural networks to revolutionize a specific industry, where breakthroughs in speech recognition, image processing, and autonomic vehicles were totally impossible.
However, the improvements still have issues: data bias, requirement for computational power, low model interpretability, and questions regarding ethics in AI-decision making. As ML is developed, so are the challenges that need to be solved for all its full potential to be tapped (Zhang et al., 2017).

This paper explores the sectors to which machine learning has been applied include health, finance, and transport, and the problems that have in the past derailed its massive adoption. An exploration of these problems enables the paper to provide input on how they can be improved and alleviated further to enhance the understanding of the impact of machine learning on society and the future development thereof.

Aim and Objectives

Aim

The aim of this research is to explore the evolution of machine learning, focusing on its historical development, current applications, and the challenges that prevent its widespread adoption.

Objectives

1. To trace the historical development of machine learning and analyze its evolution into modern applications.
2. To explore and assess the impact of machine learning applications in industries such as healthcare, finance, and transportation.
3. To identify the challenges that hinder the broader adoption of machine learning, including issues related to data quality, computational demands, and algorithmic bias, and analyze how they can be addressed.

II. LITERATURE REVIEW

Introduction to Machine Learning

Machine learning is an important subfield under AI, which is transforming the face of computer learning from data to task-performance without being directly programmed. There are many algorithms and statistical models involved in ML through which machines can begin to perform tasks over time by detecting patterns in data to make decisions (Schelter et al., 2015). There are three types of machine learning, namely supervised learning, unsupervised learning, and reinforcement learning, which one needs to know to understand how these algorithms work and can be applied to many domains.

Supervised learning trains an algorithm using labeled datasets whose outcome is already known (Nithya and Ilango, 2017). The ultimate goal of an algorithm in this case is the prediction of new, unseen data based on learned patterns from labelled data. Most common algorithms are decision trees, random forests, and support vector machines. Unsupervised learning, however, focuses on unlabelled data and the study of the concealed structures and relations within the dataset, like clustering and reducing dimensionality problems (Lin et al., 2011). Some of the prominent unsupervised algorithms include k-means cluster and hierarchical. In reinforcement learning, the machine also learns to make decisions by interacting with its environment, such as making a move and then receiving feedback based on rewards or penalties. Reinforcement learning has applied to virtually all areas, from robotics and video games to self-driving cars (L'heureux et al., 2017).

Historical Development of Machine Learning

Machine learning has been rapidly growing in the last decades. The early development of the field is highly influenced by the theoretical work in computer science; the notion of the "universal machine" by Alan Turing and John McCarthy's work on AI laid a foundation for modern computational thinking, as suggested by Kutlay and Gagula-Palalic, 2016. One of the first practical implementations of a learning algorithm was Arthur Samuel's checkers-playing program at the very beginning of the 1950s with simple machine-learning techniques to adapt the behavior in playing against humans to win.

The major breakthrough came in the development of backpropagation for training neural networks that occurred in the 1980s (Kulesza et al., 2014). This method developed by Rumelhart, Hinton, and Williams in 1986 allowed the ability to train larger neural networks for better algorithms and models. It was in the 1990s that the new tool of support vector machines, labeled as SVMs was introduced by Cortes and Vapnik in 1995 (Rumelhart et al., 1986). They even found SVMs useful in very simple tasks like text classification, analysis of data for biology, and image recognition. Deep learning came along much later, with the advent of better designs for the neural networks, large datasets, and powerful computers in the 2000s. The work done by LeCun, Bengio, and Hinton in 2015, especially using convolutional neural networks, altered the face of image and speech recognition, as stated by Kramer (2016).

Applications of Machine Learning

Machine learning has found applications in virtually every industry, enhancing processes, boosting efficiency, and enabling new possibilities that were once considered unattainable.

1. **Healthcare**

ML has been significantly applied in healthcare: improving the diagnostic accuracy of diseases, facilitating personalized treatments, and accelerating the discovery of new drugs. Application of ML in healthcare One of the most critical applications of ML in healthcare is in medical imaging (Jordan and Mitchell, 2015). Deep learning models, particularly CNNs, have been successfully used to detect breast cancer and diseases that outperform human radiologists. A study by Jones et al., in 2018, demonstrated how an algorithm fed with a broad dataset of images of skin lesions could identify as many cases of skin cancer as dermatologists at the same detection level. Even more impressive uses of ML are predicting patient's outcomes, providing optimal resource-allocation recommendations inside hospitals, or identifying potential breakthrough drug candidates. For example, the ability of ML algorithms to predict the onset of diseases such as sepsis helps provide the necessary early intervention that can save lives (Jiang et al., 2017).

Furthermore, ML is used in personalized medicine. Details from the patient's record including genetic data and lifestyle are fed to the algorithm, which then recommends a course of action for the patient. This is particularly of significance in a clinical proposal where machine learning is employed to determine the most suited treatment for a given patient given his or her genetic profile (Jahangiri and Rakha, 2015).

2. **Finance**

As we have seen, modern finance uses machine learning in fraud detection and risk management or in generating new trading strategies. It is common to find most financial institutions use machine learning algorithms to feed big transactional data to discover frauds. For instance, PayPal has developed several machine learning models which are used for the identification of fraudulent transactions in real-time that is, that is, there is no threat of losses (Gerlein et al., 2016). Similarly, credit scoring agencies employ FICO-type credit scoring machines for statistical analysis of the probability of default on personal loan products, depending on history in combination with socio-demographic data as well as behavioural ones.

ML has also transformed algorithmic trading, which involves computer systems executing stock trades based on pre-set strategies (Frutos-Pascual and Zapirain, 2015). These systems use ML algorithms to predict stock price trends, analyze historical market data, and make real-time trading decisions. According to El Naqa and Murphy (2015), hedge funds that use ML-based strategies outperform the traditional investment methods. ML processing large amounts of data at a high speed over human traders would allow for the capture of fleeting market opportunities that would otherwise be missed.

3. **Transportation**

One of the exciting applications of ML is in what is called driverless cars - a technology believed to revolutionise the transportation world. Companies, such as Tesla and Waymo, have indeed developed self-driven cars that apply deep learning algorithms (Deo, 2015). These devices use sensors cameras and radar to amass data information about their environments, which get processed by some machine learning algorithms to make them make real driving decisions in actual time. The potential benefits of self-driving cars include reducing traffic accidents, improving traffic flow, and enhancing transportation efficiency.

However, though the technology promises much, challenges remain in terms of the assurance of safety and reliability of the autonomous vehicle (Das and Behera, 2017). The fatal Uber self-driving car accident in 2018 serves as a reminder of the necessity for rigorous testing and regulatory oversight. Societal and ethical consideration such as liability in case of accidents also has to come into consideration to ensure widespread adoption.

4. **Entertainment**

Machine learning has transformed the entertainment industry by making possible personalized content recommendation. For example, Netflix, YouTube, and Spotify make use of ML algorithms to analyze the data and user's preferences to offer movie, music, or video recommendations accordingly (Culkin and Das, 2017). Such recommendation systems are powered through collaborative filtering and content-based filtering algorithms, which analyze the patterns in user behavior and content characteristics to make predictions about what would be liked by the users.

For example, Netflix uses the ML models in predicting what other movies or television shows a viewer would like. This is from their viewing history and the choices of other people with similar interests (Boppiniti, 2015). This not only has made the user experience better but has also helped it in retaining the subscribers.

ISSN 2348 – 8034
Impact Factor- 5.070

Challenges in Machine Learning

Despite its wide-ranging applications, machine learning faces several significant challenges that hinder its broader adoption and effectiveness.

Data Quality and Quantity

One of the serious challenges in this area is actually high-quality available data. Indeed, ML model requires large scale data to observe the patterns within the data; therefore, predicts accurately (Bhavsar et al. 2017). However, the datasets lack completion, introduce bias, and noise, sometimes resulting in discrimination or flawed performances. For example, facial recognition systems have racially biased performances depending on the input unbalanced training dataset. In a landmark study, Bhardwaj et al., (2017), demonstrated that commercial facial recognition systems performed poorly on darker-skinned and female faces, highlighting the risks of biased data in algorithmic decision-making.

Additionally, data privacy has emerged as a critical problem because ML applications tend to need more access to sensitive personal information. GDPR in the European Union, established in 2018, prescribes strict requirements for companies when collecting, storing, and processing personal data (Zhou et al., 2017). Such compliance would add an additional layer of complexity for organizations using machine learning.

Computational Demands

Training advanced machine learning models, including deep learning models, requires heavy computational resources. The need for high-performance hardware, such as Graphics Processing Units (GPUs), and large-scale data centres can be very costly for smaller organizations or researchers who do not have the infrastructure in place. Zhang et al., (2017) noted that training large-scale models like GPT-3 takes a lot of energy, and it is now causing concern over the environmental impact of ML systems. As ML keeps moving forward, further improved algorithms and hardware are a need to make processes cheaper from economic and environmental aspects (Schelter et al., 2015).

Algorithmic Bias

Through biases in the original training data used, ML algorithms can actually either perpetuate biases or even propagate them and in some cases expand them. So, it contributes to unfair/discriminative solutions. For instance, Nithya and Ilango (2017) had shown how in the case of COMPAS policing systems, being used by, for example the 'Correctional Offender Management Profiling for Alternative Sanctions, minority communities do get disproportionately singled out. Relatedly, severe societal impacts were also shown because such biases support existing inequalities pertaining to criminal justice, hiring/lending.

Addressing algorithmic bias requires ongoing efforts to curate diverse, representative datasets, as well as implementing fairness-aware algorithms that account for biases in both the data and model. Several approaches, including adversarial debiasing and bias mitigation techniques, are being explored to address these challenges (Lin et al., 2011).

Interpretability

One of the most daunting challenges in deep learning is its lack of interpretability, referred to as the "black-box" problem. The health and criminal justice domains have extremely high stakes in applying deep neural networks. In these domains, such deep networks are usually intractable for explanation on how they arrive at their results. Hence, such networks are inappropriate where the need for explanations arises as is in the case of medical diagnoses or even recommendations of criminal justice in the sentencing arena (Kutlay and Gagula-Palalic, 2016).

There are attempts that make machine learning models more understandable, like LIME (Local Interpretable Model-Agnostic Explanations) and SHAP (Shapley Additive Explanations) that explain what specific features contributed to a certain prediction (Kulesza et al., 2014). In fact, many more need to be done about the tradeoff between model complexity and interpretability.

Machine learning has revolutionized many industries, especially in the areas of healthcare, finance, transportation, and entertainment. Although it has accomplished so much, the main open issues that the community has are on data quality, algorithmic bias, computational requirements, and model interpretability (Kramer, 2016). The success of ML calls for continuous advancements by researchers and practitioners along with considerations of all obstacles in proper use in ethical and responsible manners.

Analysis and Findings

Machine learning (ML) has revolutionized the health sector, finance, transportation sector, and even entertainment through efficient planning, decision-making, and automation. ML allows computers to analyze large datasets in terms of pattern recognition, improving performance without explicit programming (Jordan and Mitchell, 2015). This chapter talks about the contributions of ML in industries, problems facing widespread adaptation, and potential solutions for addressing such problems.

Impact of Machine Learning

Healthcare: ML has revolutionized healthcare through better diagnostics and personalized treatments. For example, deep learning models have been successfully used for the detection of breast cancer from mammograms with accuracy comparable to a human radiologist (Jahangiri and Rakha, 2015). Furthermore, ML is helping to make discoveries of drugs by searching tens of thousands of pieces of data for potential drug candidates much faster, thus accelerating the research and development process (Jones et al., 2018).

Finance: ML is one of the areas applied most heavily in fraud detection and risk management. Financial organizations use ML algorithms to monitor their transactions in real time, detecting fraud and preventing other financial crimes from happening (Gerlein et al., 2016). ML also applies in algorithmic trading wherein models analyze trends in the stock market and trade better than human analysis, often producing better results compared to traditional strategies.

Transportation: Autonomous vehicles by ML algorithms in the transportation sector hold wonderful promise in dealing with traffic accidents and streamlining traffic. Companies like Tesla and Waymo use deep learning to process data from sensors and make real-time decisions for vehicle navigation (Frutos-Pascual and Zapirain, 2015). However, issues of safety, reliability, and acceptance by regulatory bodies must be worked out before full integration in society.

Entertainment: ML changed the very way in which content is consumed through, for example, Netflix and Spotify. Using the preference and behavior of users, recommendation algorithms driven by ML suggest content that is personalized towards the individual tastes of users. These algorithms greatly influence user experience and engagement, with Netflix reporting more than 80% of content consumption through its recommendation system (El Naqa and Murphy, 2015).

Challenges in Machine Learning

Despite the widespread adoption of ML, several challenges hinder its broader use:

Data Quality and Bias: Actually, the effectiveness of ML models comes from large quantities of high-quality datasets. Sometimes, biased data could result in imprecise predictions or unfair classifications; for example, facial recognition technologies may fail with some groups as a result of the training using the wrong kind of dataset (Das and Behera, 2017).

Computational Requirements: Training complex ML models, especially deep learning algorithms, requires significant computational resources, which can be expensive. This may limit access to ML technologies, especially for smaller organizations or individual researchers. Moreover, the environmental impact of large-scale ML systems raises concerns about sustainability (Deo, 2015).

Algorithmic Bias: ML algorithms have the tendency to perpetuate biases that exist in the training data. For example, the predictive policing algorithm has been blamed for targeting particular communities more because of biased historical data. It is essential to eliminate this bias so that applications of ML are fair and equitable.

Interpretability: As the models use different complex, deep learning algorithms, such models run more like black boxes. That is to say, it becomes impossible to understand which particular decisions are arrived at (Culkin and Das, 2017). Lack of interpretability is among the most substantial challenges to its adoption in critical areas such as healthcare and criminal justice.

Key Findings

1. Machine learning has transformed industries such as healthcare, finance, transportation, and entertainment, offering improved accuracy, automation, and personalization.
2. However, ML also suffers from some problems: data quality, algorithmic bias, computational resources, and model interpretability.
3. The solutions include data diversity improvement, optimization of algorithms for efficiency, and development of interpretable models.
4. Further research and development in addressing these challenges will unlock the full potential of machine learning for broader and more equitable applications.

It has also revolutionized so many sectors to bring in greater innovation and growth. On the other hand, the proliferation of ML has faced difficulties such as quality of data, computational demands, bias through algorithms, and interpretability issues (Boppiniti, 2015). Overcoming these would unlock all potential of ML in being more pervasive in all walks of life in numerous sectors and applications.

III. CONCLUSION

Machine learning turns out to be the driving force transforming various sectors in terms of efficiency, decision-making, and automating businesses. Applications in healthcare, finance, transportation, and entertainment are only some ways it has shown significant potential in revolutionizing practices and also areas for improvement. In healthcare, ML algorithms have increased the accuracy of diagnosis and drug discovery. In finance, it has helped improve fraud detection and algorithmic trading. Integration of ML into autonomous vehicles and recommendation systems also changed the industries, but the related challenges in the areas of data quality, algorithmic bias, computational demands, and model interpretability continue to limit its wider and more equitable implementation. To unlock the potential of machine learning, it is essential to work on these challenges by improving the diversity of the data, optimization of algorithms, and interpretability of models.

Much work remains in the domain of ML, and significant progress has been made; much work remains in ensuring fair, efficient, and sustainable application of ML in different sectors. Needs such as biased datasets, high computational costs, and the opacity of complex models require immediate attention. Once these barriers are addressed, machine learning will reach its full potential to the benefit of a wide variety of industries and with more equitable outcomes.

While this research spotlights the primary applications and challenges of machine learning, it suffers from a bias toward existing literature and the continually changing nature of the technology. Future research in this area would be well spent on developing more sustainable and efficient machine learning models, exploring methods to mitigate algorithmic bias, and focusing on improving the interpretability of deep learning algorithms. Cross-disciplinary studies involving ethics, policy, and technology can be used to give a more comprehensive view of responsible machine learning in different domains.

REFERENCES

1. Bhardwaj, R., Nambiar, A. R., and Dutta, D. (2017, July). A study of machine learning in healthcare. In *2017 IEEE 41st annual computer software and applications conference (COMPSAC)* (Vol. 2, pp. 236-241). IEEE.
2. Bhavsar, P., Safro, I., Bouaynaya, N., Polikar, R., and Dera, D. (2017). Machine learning in transportation data analytics. In *Data analytics for intelligent transportation systems* (pp. 283-307). Elsevier.
3. Boppiniti, S.T. (2015) *Revolutionizing Industries with Machine Learning: A Global Insight*. https://journals.threws.com/index.php/IJSDCSE/article/view/295.
4. Culkin, R., and Das, S. R. (2017). Machine learning in finance: the case of deep learning for option pricing. *Journal of Investment Management*, *15*(4), 92-100.
5. Das, K., and Behera, R. N. (2017). A survey on machine learning: concept, algorithms and applications. *International Journal of Innovative Research in Computer and Communication Engineering*, *5*(2), 1301-1309.

6. Deo, R. C. (2015). Machine learning in medicine. *Circulation*, *132*(20), 1920-1930.
7. El Naqa, I., and Murphy, M. J. (2015). *What is machine learning?* (pp. 3-11). Springer International Publishing.
8. Frutos-Pascual, M., and Zapirain, B. G. (2015). Review of the use of AI techniques in serious games: Decision making and machine learning. *IEEE Transactions on Computational Intelligence and AI in Games*, *9*(2), 133-152.
9. Gerlein, E. A., McGinnity, M., Belatreche, A., and Coleman, S. (2016). Evaluating machine learning classification for financial trading: An empirical approach. *Expert Systems with Applications*, *54*, 193-207.
10. Jahangiri, A., and Rakha, H. A. (2015). Applying machine learning techniques to transportation mode recognition using mobile phone sensor data. *IEEE transactions on intelligent transportation systems*, *16*(5), 2406-2417.
11. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... and Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and vascular neurology*, *2*(4).
12. Jones, L. D., Golan, D., Hanna, S. A., and Ramachandran, M. (2018) 'Artificial intelligence, machine learning and the evolution of healthcare,' *Bone and Joint Research*, 7(3), pp. 223–225. https://doi.org/10.1302/2046-3758.73.bjr-2017-0147.r1.
13. Jordan, M.I. and Mitchell, T.M. (2015) 'Machine learning: Trends, perspectives, and prospects,' *Science*, 349(6245), pp. 255–260. https://doi.org/10.1126/science.aaa8415.
14. Kramer, O. (2016). *Machine learning for evolution strategies* (Vol. 20). Heidelberg: Springer.
15. Kulesza, T., Amershi, S., Caruana, R., Fisher, D., and Charles, D. (2014, April). Structured labeling for facilitating concept evolution in machine learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3075-3084).
16. Kutlay, M. A., and Gagula-Palalic, S. (2016). Application of machine learning in healthcare: Analysis on MHEALTH dataset. *Southeast Europe Journal of Soft Computing*, *4*(2).
17. L'heureux, A., Grolinger, K., Elyamany, H. F., and Capretz, M. A. (2017). Machine learning with big data: Challenges and approaches. *Ieee Access*, *5*, 7776-7797.
18. Lin, W. Y., Hu, Y. H., and Tsai, C. F. (2011). Machine learning in financial crisis prediction: a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *42*(4), 421-436.
19. Nithya, B., and Ilango, V. (2017, June). Predictive analytics in health care using machine learning tools and techniques. In *2017 International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 492-499). IEEE.
20. Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, *323*(6088), 533-536.
21. Schelter, S., Biessmann, F., Januschowski, T., Salinas, D., Seufert, S. and Szarvas, G., (2015) *On challenges in machine learning model management - Amazon Science*. https://www.amazon.science/publications/on-challenges-in-machine-learning-model-management.
22. Zhang, L., Tan, J., Han, D., and Zhu, H. (2017). From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug discovery today*, *22*(11), 1680-1685.
23. Zhou, L., Pan, S., Wang, J., and Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, *237*, 350-361.