



Study Of Machine Learning And Big Data

Tamanna Noor M Fatakre

Department of IT

GMVCS Tala

University of Mumbai

Dhanashri Samir Vadake

Department of IT

GMVCS Tala

University of Mumbai

Nitin Namdev Pawar

Department of IT

GMVCS Tala

University of Mumbai

Prof. Avani A. Amburle

Assistant professor GMVCS University of Mumbai

Abstract :- Machine Learning is a Science Which was found and developed as a subfield of artificial intelligence in the 1950s. The first steps of machine learning go back to the 1950s but there were no significant researches and developments on this science. However, in the 1990s, the researches on this field restarted, developed and have researches to this day, it is a science that will improve more in future. In the current age of the Forth Industrial Revolution, the digital world has a wealth of data, Such as Internet of Things (IOT) data, cybersecurity data, mobile data, business data, social media data, health data, etc. To intelligently analyze these data and develop the corresponding smart and automated applications, the knowledge of artificial intelligence, particularly machine learning is the key. Machine learning is based on the principle of finding the best model for new data among the previous data thanks to increasing data. Therefore, machine learning researches will go on in parallel increasing data. There are various types of Machine Learning algorithm such as supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning exists in the area. Besides, the deep learning, which is part of a border family machine learning methods, can intelligently analyze the data on large scale. The aim of this study is to transmit the knowledge on machine learning, which has become very popular now days, and its applications to the researches.

Keywords:- Machine Learning, Deep Learning, Big Data, Artificial Intelligence.

I. INTRODUCTION

Machine learning, by its definition, is a field of computer science that evolved from studying pattern recognition and computational learning theory in artificial intelligence. It is the learning theory in artificial intelligence. It is the learning and building of algorithms that can learn from and make predictions on data sets. Since their evolution, human have been using many types of tools to accomplish various task in a simpler way. According to Arthur Samuel Machine Learning is defined as the field of study that gives computer ability to learn without being explicitly programmed. Arthur Samul was famous for his checkers playing program. Machine Learning is used to teach machine how to handle the data more efficiently. Sometimes after viewing the data, we cannot interpret the extract information from the data. In that case, we apply machine learning with the abundance of datasets available, the demand for machine learning is in rise. The purpose of machine learning is to learn from data.

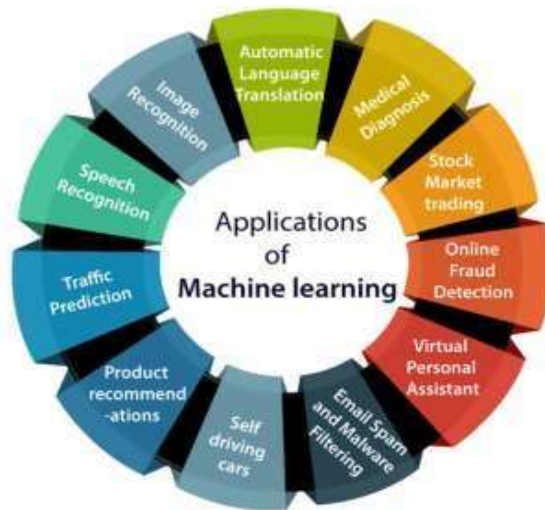
What is Machine Learning?

Machine learning is an application of Artificial Intelligence (AI) that provides system the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning is a method of approximating a function that maps input space to output space by extracting compressed no redundant information from data samples. It differs from rule-based programming in that rule-based system are based on hand-crafted rules that are composed to carry out intelligent task, while machine learning techniques attempt to formulate a function that maps input space to output space. Various types of machine learning algorithms such as supervised, unsupervised, semi-supervised, and reinforcement learning exist in the area. Besides, the deep learning, which is part of a broader family of machine learning method can intelligently analyze the data on a large scale. **Types of Machine Learning:-**

1. **Supervised Machine Learning:** Supervised learning is defined as when a model gets trained on a “Labelled Dataset”. Labelled datasets have both have input and output parameters. In Supervised Learning algorithms learn to map points between inputs and correct outputs. It has both training and validation datasets labelled.
2. **Unsupervised Machine Learning:** Unsupervised Learning is a type of machine learning techniques in which an algorithm discovers patterns and relationships using unlabelled data. Unlike supervised learning, unsupervised learning doesn't involve providing the algorithm with labelled target outputs. The primary goal of unsupervised learning is often to discover hidden patterns, similarities, or clusters within the data, which can then be used for various purposes, such as data exploration, visualization, dimensionality reduction.
3. **Semi-Supervised Learning:** Semi-Supervised learning is a machine learning algorithm that works between the supervised and unsupervised learning so it uses both labelled and unlabelled data. Its particularly useful when obtaining labelled data is costly, time-consuming, or resource-intensive. This approach is useful when the dataset is expensive and time-consuming. Semi-supervised learning is chosen when labelled data requires skills and relevant resources in order to train or learn from it.
4. **Reinforcement Machine Learning:** Reinforcement machine learning algorithm is a learning method that interacts with the environment by producing actions and discovering errors. Trial, error, and delay are the most relevant characteristics of reinforcement learning. In this technique, the model keeps on increasing its performance using Reward Feedback to learn the behavior or pattern. These algorithms are specific to a particular problem e.g. Google Self Driving cars, Alpha Go where a bot competes with humans and even itself to get better performers in Go Game. Each time we feed in data, they learn and add the data to their knowledge which is training data. So, the more it learns the better it gets trained.

Research Through Innovation

Applications of Machine Learning:



1. **Image Recognition:** Image recognition is one of the most common applications of machine learning. It is used to identify objects, persons, places, digital images, etc. The popular use case of image recognition and face detection is, Automatic friend tagging suggestion.

2. **Speech Recognition:** Speech recognition is a process of converting voice instructions into text, and it is also known as “Speech to text” or “Computer speech recognition”. At present, machine learning algorithms are widely used by various application of speech recognition. Google assistant, Siri, Cortana, and Alexa are using speech recognition technology to follow the voice instructions.

3. **Traffic prediction:** If we want to visit a new place, we take help of Google Maps, which shows us the correct path with the shortest route and predicts the traffic conditions. It predicts the traffic conditions such as whether traffic is cleared, slow moving or heavily congested.

4. **Product recommendations:** Machine learning is widely used by various e-commerce and entertainment companies such as Amazon, Netflix, etc., for product recommendation to the user. Whenever we search for some product on Amazon, then we started getting an advertisement for the same product while internet surfing on the same browser and this is because of machine learning. Google understands the user interest using various machine learning algorithms and suggests the product as per customer interest.

5. **Self-driving cars:** One of the most exciting applications of machine learning is self-driving cars. Machine learning plays a significant role in self-driving cars. Tesla, the most popular car manufacturing company is working on self-driving car. It is using unsupervised learning method to train the car models to detect people and objects while driving.

6. **Email Spam and Malware Filtering:** Whenever we receive a new email, it is filtered automatically as important, normal, and Spam. We always receive an important mail in our inbox with the important symbol and spam emails in our spam box and the technology behind this is Machine learning. Below are some spam filters used by Gmail:

- Content filter
- Header filter
- General blacklists filter
- Rules-based filters
- Permission filters

7. **Virtual Personal Assistant:** We have various virtual personal assistants such as Google assistant, Alexa, Cortana, Siri. As the name suggests, they help us in finding the information using our voice instruction. These assistants can help us in various ways just by our voice instructions such as play music, call someone, open an email, scheduling an appointment, etc. These virtual assistants use machine learning algorithms as an

important part. These assistant record our voices instructions, send it over the server on a cloud, and decode it using ML algorithms and act accordingly.

8. Online Fraud Detection: Machine learning is making our online transaction safe and secure by detecting fraud transaction. Whenever we perform some online transaction, there may be various ways that a fraudulent transaction can take place such as fake accounts, fake ids, and steal money in the middle of a transaction. So to detect this, Feed Forward Neural network helps us by checking whether it is genuine transaction or a fraud transaction.

9. Stock Marketing trading: Machine learning is widely used in stock market trading. In the stock market, there is always a risk of up and downs in shares, so for this machine learnings long short term memory neural network is used for the prediction of stock market trends. For each genuine transaction, the output is converted into some hash values, and these values become the input for the next round. For each each genuine transaction, there is a specific pattern which gets change for the fraud transaction hence, it detects it and makes our online transaction more secure.

10. Medical Diagnosis: In medical science, machine learning is used for diseases diagnoses. With this, medical technology is growing very fast and able to build 3D models that can predict the exact position of lesions in the brain. It helps in finding brain tumors' and other brain-related diseases easily.

11. Automatic Language Translation: Nowadays, if we visit anew place and we are not aware of language then it is not a problem at all, as for this also machine learning help us by converting text into our known language. Google's GNMT (Google Neural Machine Translation) provide this feature, which is a Neural Machine Learning that translates the text into our familiar language, and it called as automatic translation.

II. METHODOLOGY

- **Translation:** Translation is a natural fit for machine learning. The large amount of written material available in digital formats effectively amounts to a massive data set that can be used to create machine learning models capable of translating texts from one language to another. Known as machine translation, there are many ways that AI professionals can create models capable of translation, including through the use of rule-based, statistical, and syntax-based models as well as neural networks and hybrid approaches. Some popular examples of machine translation include Google Translation, Amazon Translate, and Microsoft Translator.

- **Chabot's:** AI-powered Chabot's can provide organizations with the additional support they need by assisting customers with most basic needs. Using Natural Language Processing, these Chabot's are capable of responding to consumers unique queries and directing them to the appropriate resources, so that customer support specialists can provide assistance to those with the trickiest of needs.

- **Generate text, images, and videos:** Generative AI is capable of quickly producing original content, such as text, images, and videos, with simple prompts. In effect, many organizations and individuals use generative AI like ChatGPT and DALL-E for a wide range of reasons, including to create web copy, design visuals, or even produce promotional videos.

- **Video Surveillance:** Video surveillance system powered AI is used to detect crime before it occurs. AI-based video surveillance system makes it possible for user to sift through various actions or objects in all the footage considerably more quickly than a human operator could. As a result, the system may inform individuals, which can eventually aid in preventing failures.

- **Virtual PAs:** Siri, Alexa, and Google Assistant are some well-known machine learning examples of virtual personal assistants. They assist you in getting almost every piece of information when asked for by voice, as the name indicates. Simply activate them and inquire,

“What is my schedule today?” or similar inquiries. “Your assistant will hunt the information to answer your inquiries, recollect them, or issue an order to other resources to gather information.

- **Statistical arbitrage:** Using an automated trading approach called arbitrage, a lot of securities are managed in the financial sector. The tactic makes use of economic data and correlations to analyze a group of securities using a trading algorithm. Examples of statistical arbitrage in the real world are- Trading algorithms that examine the microstructure of a market, analyze a lot of data, and find prospects for real- time arbitrage. The arbitrage approach is optimized by machine learning technologies to improve outcomes.
- **Automating employee access control:** Machine learning algorithms are being actively used by organizations to determine the level of access that employees would require in various locations based on their job profiles. This is one of the most intriguing machine learning applications.
- **Agriculture:** For high-quality produce, machine learning in agriculture provides accurate and effective farming with minimal labour. Additionally, Machine learning offers priceless crop-related information and suggestions so that farmers can reduce losses.

III. CONCLUSION

Machine Learning is a technique of training machines to perform the activities a human brain can do, bit faster and better than average human-being. Today we have seen that the machine can beat human champions in games such as Chess, Alpha Go, which are considered very complex. Machine Learning is concerned with computer programs that improve their performance through their experiences with their audiences. Machine learning improves the technological world with its features. The main objective of machine learning is to make the machines or devices learn the human behaviors and read their minds by studying their regular activities. Nowadays, in modern society every country follow the machine learning concept.

REFERENCES:-

- [1] W. Richert, L.P. Coelho, “Building Machine Learning System with Python”, Packet Publishing Ltd., ISBN 978-1-7826-140-0.
- [2] J. M. Keller, M. R. Gray, J.A. Givens Jr., “A Fuzzy k-Nearest Neighbour Algorithm”, IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-15, No. 4, August 1985
- [3] <https://www.geeksforgeeks.org/machine-learning>
- [4] S. Marshland, Machine Learning: an algorithmic perspective. CRC press, 2015
- [5] https://en.wikipedia.org/wiki/Instance-based_learning
- [6] P. Harrington, “Machine Learning in action”, Manning Publications Co., Shelter Island, New York, 2012