Machine Learning: Applications, Challenges and Tools

Pembarti Rachana Devidas[1], Dr. Rajesh Koolwal[2], Dr. Amjan Shaik[3]

[1]Research Scholar, Department of CSE, J.S. University, Shikohabad, Firozabad (Dist.), U.P.- 283135
[2]Assistant Professor, Department of CSE, J.S. University, Shikohabad, Firozabad (Dist.), U.P.- 283135
[3]Professor, Department of CSE, St.Peter's Engineering College, Maisammaguda, Hyderabad, T.S. - 500090

ABSTRACT

At this point in time, a great amount of data is readily available for access. As a result, it is of the utmost importance to conduct an analysis of these data in order to extract any information that may be of use, and then to construct an algorithm based on analysis. Machine learning and data mining are two methods that might be used to accomplish this goal. Machine learning is a kind of artificial intelligence that is used to construct algorithms that are based on the correlations and patterns that have been learnt from previous data. The field of machine learning has a wide range of subfields. The many applications of machine learning, its needs, its methodologies, the various approaches it takes to problem resolution, and the challenges it faces are all topics that are covered in this article.

Keywords: **Machine learning, data mining, artificial intelligence.**

INTRODUCTION

Over the last several decades, artificial intelligence (AI) has developed into a vast and intriguing topic within the discipline of computer science. This is due to the fact that it allows robots to carry out tasks that humans are capable of. As an additional objective, it aims to teach computers how to handle problems that occur in the actual world as effectively as is possible. As a result of advancements in science and technology, artificial intelligence (AI) systems are now able to acquire new knowledge and improve themselves based on their previous experiences without the need for a specialised support code. This is made possible by the fact that they are exposed to new data. The end consequence of this is the development of machine learning (ML) technology, which employs learning algorithms to acquire knowledge from the data that is readily accessible [1,2].

1. MACHINE LEARNING

The use of machine learning (ML), which is a subset of artificial intelligence (AI), enables computer systems to improve the accuracy of their predictions without being specifically taught to do so [6]. For the purpose of predicting new output values, machine learning algorithms make use of historical data as input materials. The use of machine learning is often included into recommendation engines. The identification of malware threats, the detection of fraudulent activity, the filtering of spam, business process automation (BPA), and predictive maintenance are applications that are used often.
Training computers to learn on their own is the subject of the field of research known as machine learning. The problem at hand is the question of why we would want robots to be able to learn on their own. The ability of robots to do tasks that are repetitive, high-frequency, and need a high level of precision while staying bored is one of the many potential advantages of this technology. Take, for example, the responsibility of cleaning and washing off the floor. It is dependent on the person who is doing the labour as to whether or not the output is consistent. After working for a few hours, we experience feelings of exhaustion or boredom, and the possibility of being ill has an impact on the result.

On the other hand, if we could teach machines to detect when a floor needs to be cleaned and mopped, as well as how much cleaning is necessary depending on the kind of floor and its state, then their performance would be much improved. They will not get exhausted or unwell, and they will continue to work!

The purpose of machine learning is to accomplish this objective by providing computers with the ability to learn on their own. For the purpose of responding to questions such as

- It is necessary to sweep and clean the floor, is that correct??
- How much time should be spent cleaning the floor??

Machine learning models are useful because they provide a means for machines to reason, which is a need for machines [7]. The operation of machine learning is shown in Figure 1. It is the machines that are responsible for gathering data from the world and then feeding it into the machine learning model. In order to make predictions about things like, the model makes use of this data:
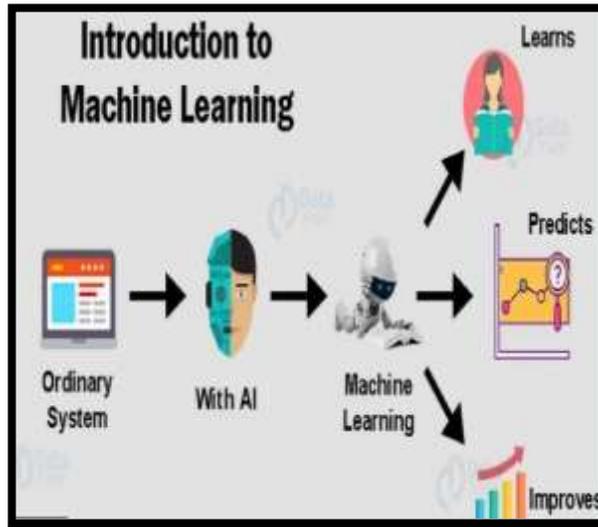

Figure1: Introduction of Machine Learning

LITERATURE REVIEW

Mariette Rahul Khanna and Awad Khanna 15 is the year. To begin, The objective of this research is to shed light on the importance of preprocessing as well as the numerous preprocessing strategies that may be used in real-world scenarios. Within the scope of this research, the various data preparation activities that are responsible for imparting meaning to raw data and preparing it for future processing are investigated.

2018 article by Teng Xiuyi1 and Gong Yuxia1 a pair of The relevance of machine learning in data mining is the primary emphasis of this paper, which also examines its definition, model, development stage, categorisation, and commercial use. In light of this, the purpose of this article is to offer an overview and analysis of machine learning technology, specifically focussing on the benefits and drawbacks associated with its use in data mining.

(2017) publication by Kajaree Das and Rabi Narayan Behera [This is the fourth point] In addition to describing some of the most regularly used machine learning algorithms, this article discusses the concept of machine learning as well as its evolution. Additionally, it makes an attempt to compare three of the most often used algorithms by using certain fundamental concepts. Using the Sentiment140 dataset, the performance of each algorithm was monitored and compared in terms of the amount of time it took to train, the amount of time it took to forecast, and the accuracy of the predictions.

5 Kotsiantis's work from 2007 A number of different approaches to categorising data using supervised machine learning are discussed in this article. It should come as no surprise that a single chapter cannot possibly cover all of the supervised machine learning classification approaches from start to finish.

Pradeep Kumar Bhatia, Yogesh Singh, and Omprakash Sangwan, in addition to Omprakash Sangwan, Yogesh Singh, Pradeep Kumar Bhatia, and Omprakash Sangwan An in-depth analysis of the research study on expert software creation estimates via the use of machine learning techniques (MLT) is presented in this article. In the present day, machine learning is demonstrating that it is capable of providing projections that are accurate on a constant basis. Neural networks, case-based reasoning, classification and regression trees, rule induction, genetic algorithms, and genetic programming are the most common machine learning approaches for expert estimates in the software development industry. Other prominent methods include genetic programming and genetic algorithms.

In every one of our studies, we discovered that the outcomes of different machine-learning strategies differ depending on the application domains to which they are applied. All of these variations were observed. The study of the research that was conducted by the authors reveals that these approaches are sensitive to the data that they are trained on and that they are competitive with conventional estimators when applied to a single data set by the authors.

2. APPLICATIONS OF MACHINE LEARNING IN REAL LIFE

After gaining an understanding of it, you may be interested in learning more about the applications of machine learning [2,] as well as how it influences our day-to-day lives. With the exception of those who have been living under a rock, you have already been subjected to a considerable influence brought about by machine learning.

Let's have a look at a few different types of circumstances in which we have previously used the results of machine learning.

• The use of Facebook, LinkedIn, or any other social media site to promote your friends and advertise products or services that you may be interested in
• The products that Amazon recommends to you based on your previous searches practicable.
• Institutions that use machine learning to identify fraudulent activity in real time

➢ **Need of Machine Learning**
The volume of data that is being created is rapidly increasing as a result of the decrease in the cost of sensors (Force 1)[7]. Figure 2 shows a representation of the data's degree of variation.Every time you do an activity on any website, such as Facebook or YouTube, you are doing something that generates data for these corporations. Every single linked gadget, such as fitness bands, smart watches, and other connected pieces of equipment, generates data.

It is now much more affordable to store this information, which is a significant factor (Force 2).

Considerable progress has been made in reducing the cost of computing (Force 3).

Computing has become more accessible to the general people as a result of cloud computing (the fourth force).

A world in which we are able to not only produce more data but also store it at a low cost and execute large calculations on it is created as a result of the convergence of these four factors. Despite the fact that the methods and algorithms of machine learning were well-known at the time, this was not
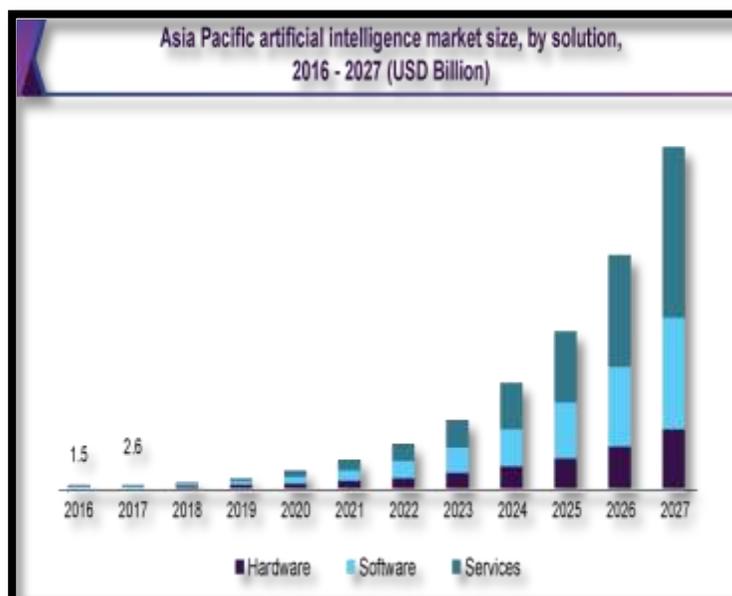

Figure 2: Data Variation Graph

➢ **Tools Used In Machine Learning**

The field of machine learning employs a wide range of methodologies and programming languages [9]. The scale of your activities and the requirements you have will define the tool that you choose to use. On the other hand, the following are the machine learning tools that are used the most frequently:

Access to the programming language used for machine learning may be gained via the use of machine learning software. Best practices for both methods and applications are provided by them [8]. There are frameworks that are provided by machine learning tools, which make it possible to carry out a project or module. The following are some examples of platforms that are used for machine learning that:

- Examples of SciPy subparts for Python are scikit-learn and Panda.
- R Platform.
- WEKA Machine Learning Workbench.

There are several libraries that are included in machine learning tools. These libraries provide all of the capabilities necessary to finish a project, and within these libraries, there are many different algorithms. The majority of libraries are:

- JSAT in Java.
- scikit-learn in Python
- Accord Framework in .NET

Some of Databases used by machine learning are:

- SQL
- Oracle
- Hadoop

➢ **Problems Solved By Machine Learning**

Machine Learning problems can be divided into 3 broad classes:
- Supervised Machine Learning
- Classification Problems
- Regression Problem
- Unsupervised Machine Learning
- Reinforcement Learning

1. Supervised Machine Learning:
When you have historical data containing outcomes (also known as labels in the vocabulary of machine learning) and you want to predict prospective outcomes, you may use algorithms that are supervised machine learning [3, 5]. It is possible to categorise problems that include supervised machine learning into two distinct categories:

(i) Classification Problems: When it is necessary to divide the findings into a number of distinct categories. For example, determining whether the floor needs to be mopped or cleaned is an example of a categorisation difficulty. The outcome will be classified as either yes or no, depending on the circumstances. In a similar vein, evaluating whether or not a client will fail on their loan is a categorisation problem that each and every bank is interested in finding a solution to.

(ii) Regression Problem: There are problems with regression that arise if you are attempting to determine the cost of something. As an instance, assessing the amount of cleaning that is necessary is an example of a regression problem. When the quantity of default that is projected for a client is uncertain, this is an example of a regression problem.

2. Unsupervised Machine Learning:
It is not always desirable to have the ability to accurately forecast the outcome of a situation [4,10]. All that is required of you is to do some segmentation or clustering. To better understand the activities of its consumers, for instance, a bank would seek to divide its clientele into several categories. This is an example of an unsupervised machine learning problem since we are not making any predictions about the results.

3. Reinforcement Learning:

On the other hand, reinforcement learning is considered to be the best chance for genuine artificial intelligence. And this is precisely why, given the huge potential that Reinforcement Learning has. In comparison to traditional machine learning, this topic is more difficult to understand, yet it will be just as significant in the years to come. If you are looking for an introduction to reinforcement learning, this article is as good as any other you could have found.
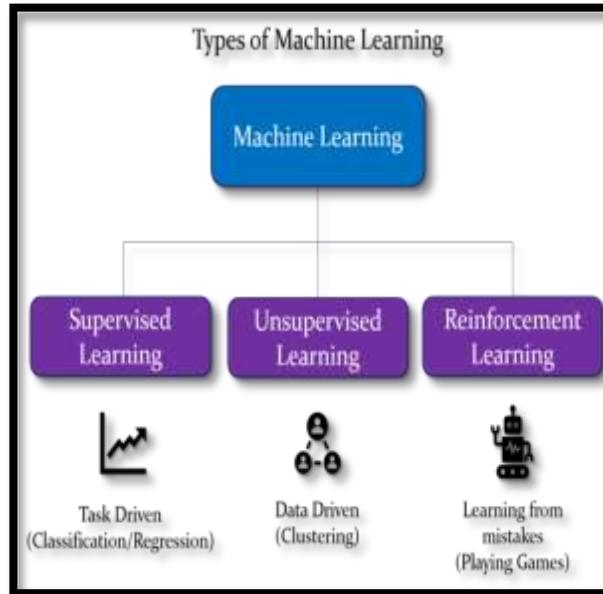

Figure 3: Types of Machine Learning

Above figure depicts the types of ML.

➤ **Challenges of Machine Learning**

There are still a number of substantial obstacles to overcome in the field of machine learning, despite the fact that it has made great progress in recent years [12]. Challenges in machine learning are shown in Figure 4. It is a topic that is actively being researched, and we anticipate a significant amount of work to be put forward in order to find solutions to these issues in the not too distant future.

1. Massive amounts of data are required: Today, training a model necessitates a massive amount of data.

2. Needs a lot of computing power: In the present day, machine learning and deep learning models need a significant amount of processing power in order to do basic tasks (tasks that humans consider to be easy). Because of this, the use of specialist technology, such as graphics processing units (GPUs) and teraflops (TPUs), is required. It is necessary to work on lowering the cost of computations in order for machine learning to make a meaningful impact..

3. Model analysis can be complicated at times: There are certain modelling techniques that come with a great degree of accuracy yet are difficult to explain. The owners of businesses may find this to be quite frustrating. Imagine you are a banker and you are unable to provide an explanation for why a request for a loan from a client was denied!

4. Need for new and updated algorithms: When it comes to finding solutions to some of the problems that were described above, researchers are still looking for new and better algorithms.

5. There is a need for more data scientists: In addition, since the field has grown at such a rapid pace, there are not a lot of individuals who possess the necessary skill sets to tackle the diverse range of issues. This is something that is anticipated to continue in the years to come. Therefore, if you are thinking of pursuing a career in machine learning, you are not alone in your decision!

Figure 4: Challenges in Machine Learning

CONCLUSIONS

The different applications of machine learning, the requirements of machine learning, the numerous methods used by machine learning, the various sorts of approaches to problem resolution, and the obstacles that machine learning encounters are all topics that are covered in this article.

Within this study, the many approaches to machine learning have been dissected and examined. This report also included a discussion of the many types of machine learning approaches that are now available. There are many different processing tools and machine learning implementations that are currently being researched and experimented with. In the part on the review of the relevant literature, we investigate how various machine learning algorithms have fared over the course of the last several years in a variety of domains when integrated with traditional methods, as well as how these algorithms have outperformed earlier models.

REFERENCES

[1]. Mariette Awad, Rahul Khanna. "Efficient Learning machines: Concepts and Applications". *Aspress Publishers,* 2015.
[2]. Teng Xiuyi1,Gong Yuxia1. "Research on Application of Machine Learning in Data Mining". IOP Conf. Series: Materials Science and Engineering, 2018.
[3]. M. Praveena,V. Jaiganesh, "Literature Review on Supervised Machine Learning Algorithms and Boosting Process". *International Journal of Computer Applications*, ISSN No. 0975 – 8887, vol. 169, 2017.
[4]. Kajaree Das, Rabi Narayan Behera. "A Survey on Machine learning: Concept, Algorithms and Applications", International Journal of Innovative Research in Computer and communication Engineering . vol. 5, 2017.
[5]. Kotsiantis."Supervised Machine Learning: A Review of Classification Techniques", Informatica. pp 249-268, 2007.
[6]. Yogesh Singh, Pradeep Kumar Bhatia & Omprakash Sangwan "A review of studies in machine learning technique". International Journal of Computer Science and Security, vol.1, pp 70 – 84, 2007.
[7]. Petersp, "The Need for Machine Learning is Everywhere" March 10, 2015.
[8]. Jason Brownlee, "A Tour of Machine Learning algorithms" November 25, 2013.
[9]. Jason Brownlee, "Machine Learning Tools", December 28, 2015.
[10]. Taiwo Ayodele, "Types of Machine Learning algorithms", 2010.
[11]. Teng Xiuyi, Gong Yuxia, Research on Application of Machine Learning in Data Mining, IOP Conf. Series: Materials Science and Engineering 392 (2018).