

FUTURISTIC ROLE OF MACHINE LEARNING: EXPLORING DOMAINS

Meenakshi Gupta

E-Mail Id: meenakshi78gupta@gmail.com

School of Engineering & Technology, Ansal University, Gurgaon, Haryana, India

Abstract-The current virtual world and the technical trends in socio - economic scenario have alluded the way for a future which calls for collaborative work between machine intelligence, huge computations, giant networks, and big data. The last decade has seen the effective application of different machine learning techniques such as unsupervised, supervised, and reinforcement techniques been used to solve various real life multidisciplinary problems. The objective of this work is to present a systematic overview on the use of data analytics and prediction techniques in various domains of industry. It discusses the roadmap of machine learning towards its potential use in various application domains. vis a vis agronomics, ecommerce, training and various other industry driven application areas. This review paper will discuss the taxonomy of machine learning including the definition and nomenclature, also its possible applications to various industry driven areas is explored. Also, various research gaps have been pointed out which can help the aspiring scholars to pursue their research in this field.

Keywords: Machine Learning, Supervised Learning, Unsupervised Learning, Algorithms, Agronomics, Ecommerce, Training Industry, prediction problems

1. INTRODUCTION

Machine learning is making its way into all kinds of products, systems, spaces, and experiences. To understand the impact in the present generic context and to utilize the potential of machine learning, we need to train a new generation of designers and thinkers. This new wave of technical paradigm shift invites a mix of thinkers, designers, data analytics, engineers, people from diverse professions to come together and collaborate towards realization of this multidisciplinary environment.

Learning is the basic means to obtain knowledge. As it is a generic process, it happens naturally to humans in the form of experiences as compared to machines which rely on data for their learning. So basically machine learning (ML) can be understood as a category of artificial intelligence which can make computing machines to learn and think on their own. It encourages PCs to alter their activities through feedback info and afterward to improve the activities to accomplish progressively precise result in the long run. Precision can be estimated as the occasions the picked activities result into right results.

The definition of Machine Learning is present across vast literature. Formally, as defined by Tom Mitchell [1], “A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E”. Although, the term finds its first mention in 1959 by Arthur Samuel [2] who defined Machine Learning as a field of study that gives learning ability to PCs without being unequivocally customized.

Machine learning is a multi-disciplinary innovation that causes PCs to get into a self-learning mode. At the point when new information is taken care of, these PCs learn, develop, change, and create without anyone else and can give predictions that can guide better decisions and smart actions in real-time without human intervention These are illustrated in the following fig. 1.1.

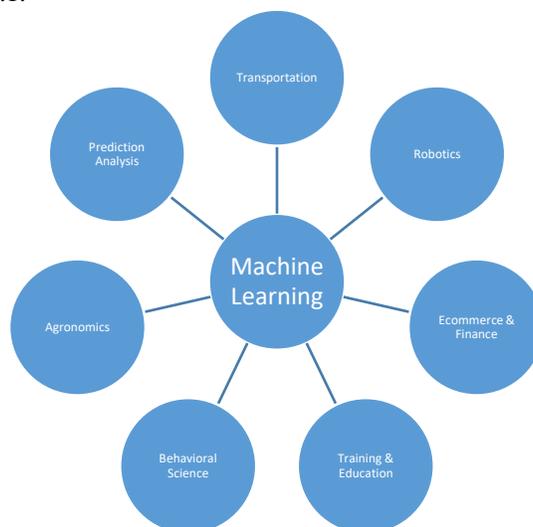


Fig. 1.1 Disciplines of Machine Learning

Genuine issues of fluctuating multifaceted nature can make brilliant contender for Machine Learning applications. It can be applied to areas of day to day activities like Google's self driving car, recommendations given by online ecommerce sites like Amazon and movies and program guides like Netflix, TVF etc, customized auto-spell feature on Whatsapp application and even detecting and predicting cyber fraud activities. Owing to affordable data storage, powerful computing capabilities, the variety and volume of big data has increased manifolds. This has prepared us to structure and program express calculations with elite yield, for instance, email spam separating, misrepresented identification on social networking sites, friend suggestion on Instagram and Facebook, online stock trading, feature detection, healthcare diagnostics, prediction analysis of traffic, stock exchange, poll results, etc, character recognition and product recommendation to name a few prevalent ones.

The primary aim of this manuscript is to give the reader an overview and idea of developing the Machine Learning journey to its present day status quo, brief introduction to various machine learning paradigms, algorithms, the most useful application domains of machine learning along with the challenges associated with them.

The paper is organized as follows: Section 2 explains the need of Machine Learning approach as alluded by various problems in data science. The roadmap of Machine Learning to its present day form is covered in Section 3 under the heading Machine Learning Development. The Paradigm of Machine Learning and its Algorithms are briefly introduced in Section 4. The role and performance of machine learning in different fields is discussed with supported literature review in Section 5. Section 6 gives the analysis and discussion. Section 7 concludes the research work.

2. NEED OF MACHINE LEARNING

New technologies of PC and pattern Recognition paved way for Machine learning as it was theorized that PCs can learn without being customized to perform explicit errands without any human supervision and intervention. The computers could learn from previous computations to produce reliable, repeatable decisions and give meaningful results. There are basically two kinds of jobs that machine learning can cater to:

2.1 Human Beings Assisted Jobs

These jobs are undertaken by human beings regularly but finish to perfection and timelines and other such constraints are there which call for the need of machine learning approach. eg mopping the floor, cleaning buildings, driving, speech to text conversion, recognition of voice etc

2.2 Jobs not Requiring Human Assistance

Jobs such as weather forecasting, availability of tickets in a particular airline, e-commerce, gathering satellite data, predicting enemy ship trajectory on the sea, etc are some examples which can be done by machine learning programs. The underlying problem of data science is volume and complexity. But with emergence of data on cloud and supercomputers, Machine learning has demonstrated capacities to innately take care of these issues. According to Hayashi and Chikio [3] information science is a primary thought through which measurements and investigation of information can be joined, and by further applying methods of machine learning, the actual phenomena can be understood and analyzed.

3. DEVELOPMENT OF MACHINE LEARNING

Machine learning has enabled systems to scan and scrutinize data and produce knowledge from it. Over time and with experience, machine learning helps in applying and improving knowledge and not just simply learn or extract knowledge. This has initiated a paradigm shift in the traditional programming. where programs are written to automate tasks. ML now can "create" the program according to the "best fit" model of data. ML techniques have now become a catalyst in various systems that handle analysis, decisions and automation ranging from home to vehicles.

Computer engineers, scientists, research scholars, students and industry professionals have explored machine language for more than 7 decades. The table 3.1 below illustrates the timeline evolution and journey of machine learning through various applications which have changed the way we humans perceive about technology.

Table-3.1 Developments in Machine Learning Across Various Domains

Year	ML Technical Development
1950	Alan Turing proposes the Turing Test as a measure of machine intelligence [5]
1951	The main working AI programs were written in 1951 to run on the Ferranti Mark 1 machine of the University of Manchester
1952	Arthur Samuel (IBM) wrote the first game-playing program [6]
1956	term "Artificial Intelligence" was first coined for the IBM conference [7]
1959	The first Artificial Intelligence Lab at MIT is formed
1960	Bayesian Methods for prediction and inductive inference were proposed by Ray Solomonoff
1966	Donald Michie and others [8] introduced Machine Intelligence at a workshop in Edinburgh
1971	A fully automatic, logic-based theorem prover called Boyer Moore theorem prover was proposed [9]

1982	A recurrent Neural Network was proposed by John Hopfield [10]
1985	A program that figures out how to articulate words a similar way an infant does, is created by Terry Sejnowski [11]
1987	a theoretical description of the mind as a collection of cooperating agents by Marvin Minsky [12]
1990s	Paradigm shift to Data driven approach programming. Programs are created to assess enormous data and "learn" from the outcomes.
1997	IBM Deep Blue supercomputer beat the world champion at chess
1998	Tim Berners-Lee published his Semantic Web Road map paper [13]
2002	software library called TORCH for machine learning is reported[14].
2006	"deep learning" term is introduced
2010	people can interact with the computer via movements and hand gestures with the Microsoft "KINECT"
2011	GOOGLE BRAIN is developed and it can simulate the brain of a cat
2012	X lab of Google uses machine learning to identify youtube videos containing cats
2014	Face book researchers release "Deep Face", a system that uses neural networks to identify faces with utmost over 97 percent correctness.
2015	Amazon releases its machine learning platform to give suggestions to its users based on their past purchases and searches
2016	Machine Learning and AI program AlphaGo developed by Google beats human champion by 6-0 in he world's most complex board-game GO
2017	Poker which is an Imperfect information Game , is palyed and won by AI and machine learning program called Deepstack [15]
2018	Google assistant DUPLEX is announced as an AI assistant to book appointments over phones and its voice quality matches that of humans
2019	Google launched its cloud based Gaming technology STADIA
2020	IIT Kharagpur launch is AI system to read legal judgments by using a more evolved machine learning technique.[16]

4. MACHINE LEARNING: COMPONENTS, PARADIGMS AND ALGORITHMS

Machine learning is an iterative and complex process by nature. Its complexity can be attributed to the complexity and volume of the data involved. From this data predictive and meaningful pattern is to be found out for a meaningful output.

4.1 Machine Learning Components

The process model of machine learning can be understood by its six modules. These are not dependent on the type of algorithm chosen. The following figure 4.1 shows the primary modules of the ML Process.

Every component of this model is accomplished as a specific task as described in the following steps

4.1.1 Data Collection

We have to carefully select and get the data that we want to work with. We should work with the domain experts in the problem area as they are people who know a lot about actual problems that we need to solve. These individuals, being specialists, realize that what information or information model that we get from the procedure is prescient. In any case, since the information with which we start is crude and unstructured, it is never in the right structure varying for genuine preparation. It could have copy of information, or information that is missing, or, more than likely a ton of additional information that isn't required. The information could be framed from different sources which may likewise in the end wind up being copy or repetitive information. This progression is significant in light of the fact that the quality and amount of information that we accumulate will straightforwardly decide how great our prescient model can be.

4.1.2 Preparation of Data

Now after data collection comes the necessity for pre-preparing the information, with the goal that the procedure could comprehend the information. So as to apply the genuine calculation to the information, we have to have that total unstructured information into an organized and molded information. This procedure is iterative until the information is accessible. In ML the significant piece of time is spent in this procedure. That is, taking a shot at the information to make it organized, spotless, prepared and accessible for additional handling.

4.1.3 Feature Selection

Feature selection refers to the process of applying statistical tests to inputs, given a specified output. The goal is to determine which features are more predictive of the output. After we get the data from above steps, we select and

chose the data relevant to our machine learning model. The left over irrelevant data can be called as noise and therefore removed from our further model.

4.1.4 Selection of Algorithm

Tradeoffs between the following justify the selection of an algorithm:

4.1.4.1 Number of Features

A feature is a quantifiable data of the process which we are analyzing for machine learning. For certain types of data, the number of features can be very large compared to the number of data points and that can make training time impractically long.

4.1.4.2 Training Time

Training means using historical data to build a machine learning model that minimizes errors. The number of minutes or hours necessary to train a model varies a great deal between algorithms. some algorithms are more sensitive to the number of data points than others

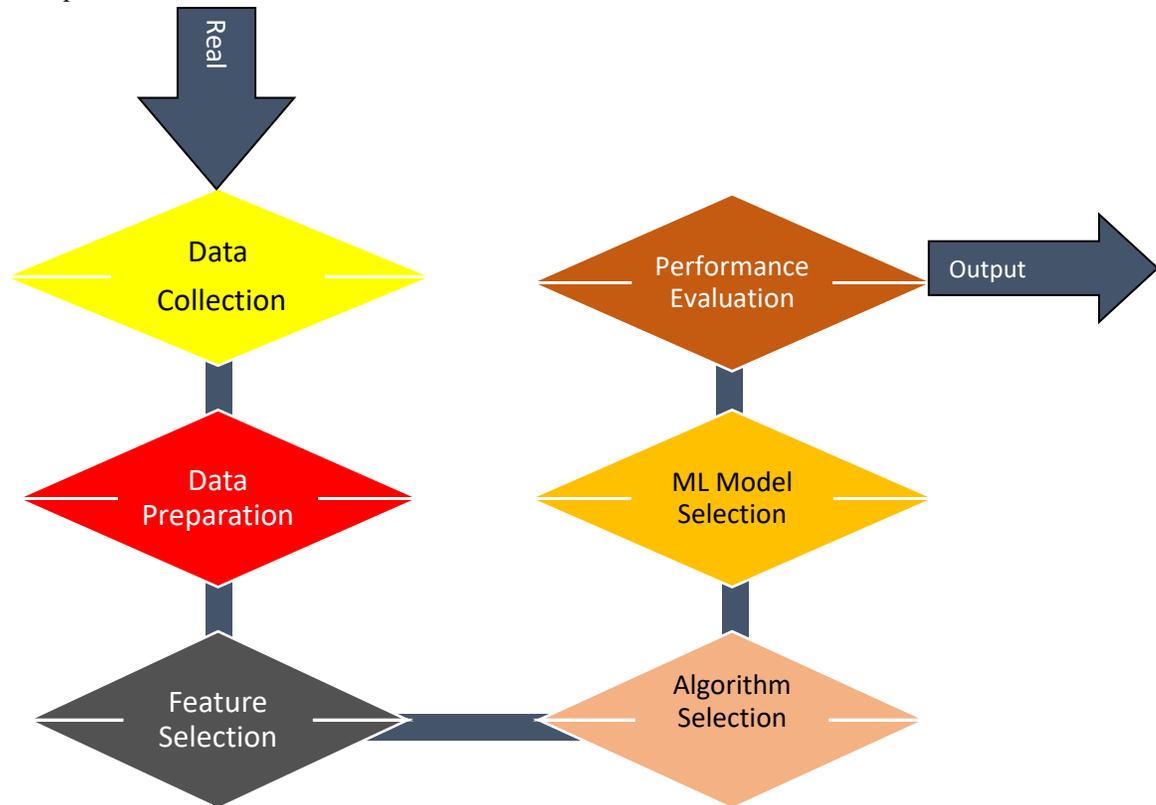


Fig. 4.1 Steps to Model a Machine Learning Process

4.1.4.3 Accuracy

Accuracy can be measured in terms of effectiveness of a model as the proportion of true results to total cases. Getting the most accurate answer possible isn't always possible. Sometimes an approximation is also enough, depending on our application program. This can also result in reduction of processing time dramatically. so we can chose more approximate methods.

As is imperative from the aforesaid discussion, selection of the best ML algorithm for the problem is crucial to get the best outcome from the ML model. A brief discussion on types of machine learning algorithms follows in Section The various ML algorithms are discussed in detail with their comparisons and tradeoffs in Section 5.

4.1.5 Selection of Model and Training

Model choice is the way toward choosing one last ML model from among an assortment of applicant models for a preparation dataset. A model that meets the necessities of the partners given the pre-designated time and assets can be chosen for additional preparation. The chosen model now should be prepared utilizing prepared information.

4.1.6 Evaluation of Performance

There are 3 metrics based on Confusion-based matrix that are used to evaluate a classification model namely accuracy, precision, and recall [16].

Accuracy is characterized as the level of right forecasts for the test information. It very well may be determined effectively by isolating the quantity of right expectations by the quantity of complete forecasts.

$$\text{Accuracy} = \text{Correct Predictions} / \text{All Predictions}$$

Precision is characterized as the portion of significant models (genuine positives) among all occurrences of the models which were anticipated to have a place in a specific class

$$\text{Precision} = \text{True Positives} / (\text{True Positives} + \text{False positives})$$

Recall is characterized as the portion of models which were anticipated to have a place with a class as for the entirety of the models that genuinely have a place in the class.

$$\text{Recall} = \text{True Positives} / (\text{True Positives} + \text{False Negatives})$$

Before constant execution of the framework, the model must be tried against inconspicuous information to assess what amount has been realized utilizing different execution boundaries like those discussed [17,18].

4.2 Machine Learning Paradigms and Algorithms

Therefore, the problem in consideration must be first categorized appropriately before going for actual problem solving. Then the machine learning approach can be explored for. A problem in data science can belong to any of the following five categories. A machine learning approach can be applied for correctly, after deciding the category of the data in consideration. The explanation for these five categories follows:

4.2.1 Classification Problem

A data mining approach which allows items in a collection to target categories or classes is known as classification (Han & Kamber, [4]) A type of problem in which the resultant output is a category like 'spam/not spam', 'Present/Absent', 'red/blue', 'disease/ no disease', '0/1' etc. It can further be of two types - a two way that is binary or multi way classification problem, depending on the number of output classes. Classification either predicts categorical class labels or classifies data (construct a model) based on the training set and the values (class labels) in classifying attributes and uses it in classifying new data.

4.2.2 Regression Problem

Regression algorithms are used to predict the outcome of an event based on the relationship between variables obtained from the dataset. They are used when the output Variable is real value or continuous. It deals with data like "salary" or "weight" and can answer questions like 'how much' or 'how many'.

4.2.3 Anomaly Detection Problem

In these types of problems, a typical pattern is detected along with some peculiar noticeable changes or anomalies in the pattern. Anomaly detection refers to identification of items or events that do not conform to an expected pattern or to other items in a dataset that are usually undetectable by a human expert. Such anomalies can usually be translated into problems such as structural defects, errors or frauds. For example, multiple failed login attempts indicating a probable cyber security threat. Any aberrations from normal behavior indicate a presence of intended or unintended induced attacks, defects, faults etc and call for attention from concerned authorities thus safeguarding the interest of the clients.

4.2.4 Clustering Problem

Clustering algorithms come under the umbrella of unsupervised learning algorithms. In this method, similar groups of data are identified in a dataset to form small group called as class or cluster and henceforth the term clustering. These different classes or clusters are then labeled. The training algorithm puts new unseen data in one of the clusters.

4.2.5 Reinforcement Problem

The concept of reinforcement problems are used when it is to be decided for whether a decision is to be made based on past experiences of learning. The machine agent employs a trial and error approach to interact with the continuously changing environment and to come up with a solution. The concept of rewards and penalties is used to get the machine agent to do what the programmer wants, without specifying any hints or suggestions on how the task is to be accomplished. Some example applications are gaming programs and pressure/temperature controlled programs.

Contingent upon how a calculation is being prepared and based on accessibility of the yield while preparing, ML ideal models can be grouped into the accompanying sub-segments.

4.2.6 Supervised Learning

It assists machine learn exclusively in learning and has data with well explained output. In this feedback is also there. The outcome result is in terms of future prediction. It is applicable for Classification and Regression Problems. Some prevalent Supervised Algorithms are:

- Decision Trees
- Naive Bayes
- SVM for classification problems
- Random Forest for both classification and regression problems
- Linear Regression for Regression problems

- Ordinary Least Square Regression
- Logistic Regression

4.2.7 Unsupervised Learning

In unsupervised learning the PC interprets data by recognizing unstructured existing patterns and therefore the evaluation is subjective and indirect. Since the training is not labeled, it does not find anything specific but can hint at various results. The main aim of this learning is to know and understand more about the input data. Unsupervised learning can be categorized further into problems of clustering and association. To name a few unsupervised Algorithms:

- K-means for clustering algorithms
- Apriori algorithm for association
- Singular Value Decomposition
- Independent Component Analysis

4.2.8 Reinforcement Learning

Reinforcement learning can be considered as an AI approach and is intermediate type of learning. It is reward based approach and works on the positive outcome of result. With time and training, the machine adapts and learns to survive in a typical environment. The whole purpose of the reinforcement algorithm is to maximize rewards. Some of the reinforcement Algorithms are:

- Q-learning
- SARSA
- Deep Q Network

5. ROLE OF MACHINE LEARNING IN VARIOUS RESEARCH AREAS

As highlighted in section 2 and 3 above, due to the availability of huge amount of data, new domains are emerging and therefore the new is becoming old at a very fast rate. So some recent advances in relevant domains are discussed and highlighted.

5.1 Role of Machine Learning in Agronomics

Amount of water required, nutrients and fertilizers usage, response towards climatic change, resistance to plant diseases, quality of taste etc are some of the parameters that can be analyzed with the help of available field data and application of Machine learning. Prediction Models to predict the availability of fittest genes with the most beneficial traits can be built. To reduce risks for farmers ML enabled applications can be built to help them predict their harvest yields as proposed by Farzaneh Sajedi-Hosseini et al. [20], rainfalls, weeds infestation and management.

5.1.1 Water Management

Choubin, Bahram et al [19] have given a novell machine learning approach to model the suspended sediment in the river which plays an important role in managing the water resources. For predicting pollution of ground water levels Farzaneh Sajedi-Hosseini et al. [20] have used Support Vector Machine (SVM), Boosted Regression Trees (BRT) and Ensemble. They even assessed the amount of fertilizers containing nitrogen that can improve the quality of groundwater.

5.1.2 Crop Yields

To predict the yield of maize Avat Shekoofa et al. [21] proposed CART method as best amongst various clustering and decision tree methods. In another study, X.E.Pantazi et al.[22] have suggested three Self Organizing Map models Counter-propagation Artificial Neural Network (CPANN), Supervised Kohonen Network (SKN) and XY-fusion network(XYF) to predict the yield of wheat grain.

5.1.3 Drought prediction

ANN and Support Vector regression with boosting techniques were employed by Belayneh et al. [23] to predict the drought in the area.

5.1.4 Crop Quality and Disease Detection

ML approach can be used to predict the quality of the crop by various techniques. One such ML techniques is being proposed by Nashwa El-Bendary et al. [24] to classify the ripeness of tomatoes at various stages of its growth thus predicting the quality and helping early disease detection.

5.2 Role of Machine Learning in Ecommerce

Ecommerce sector generates vast amount of data that can be used by machine learning programs to find structured patterns which can play important role in future trends like prediction of stocks, investment suggestions, prediction of entertainment industry like performance of movies, prediction of winners in games like racing, risk analysis for banks, operation analysis for various industries etc. All these advancements in business world greatly influences the economy of a country and therefore the emphasis on machine learning is ample.

5.2.1 Customer Prediction

Junichiro Mori et al.[25] and T. Vafeiadis et al.[26] have employed SVM and DT(Decision Tree) using Back Propagation Networks and other ML methods to predict customers for a company

5.2.2 Prediction in Sports

According to Bigus[26], using neural systems, a piece of computerized reasoning for the forecast of the result of sports like rugby and football. Multilayer perceptrons were utilized in anticipating the outcome. Utilizing multilayer perceptrons give an extra preferred position to neural systems. This is by all accounts productive strategy since the multilayer perceptron adjusts rapidly and performs well in forecast. Essentially, this strategy can be utilized to other betting games like pony dashing and horse racing.

5.2.3 Prediction of Stock Market

Yuqing He et al. [27] gave twelve stock market trend indicators for analysis. Authors have also studied the three different feature selection algorithms viz. Principle Component Analysis (PCA), Genetic Algorithm (GA) and Sequential Forward Selection (SFS) along with their merits and de-merits.

5.2.4 Bankruptcy Prediction

Nanxi Wang [28] have given a machine learning model using SVM and Neural Network to predict about Bankruptcy which can play a very vital role in the operational and risk management of a banking organization.

5.3 Role of Machine Learning in Multidisciplinary Domains

The traditional methods of various problem domains like software engineering, chemical industry, architecture, hardware machine industry can be replaced by viable alternatives of Machine learning.

5.3.1 Software engineering

Oliveira et al. [28] have worked on software effort and software accuracy, two important metrics in software reliability. The authors used SVM and Decision Trees for their model. To help enhance and evaluate the software quality Malhotra et al. [29] have proposed their ML method which uses Ensemble techniques of Random Forest along with SVM and four other machine learning techniques.

5.3.2 Manufacturing and Mechanical Industry

Early detection of faults in Bearings is very crucial as it a vital element in manufacturing and household devices. Jagath Sri Lal Senanayaka et al. [30] have given classification of faults in bearings along with prediction of faults in them. The authors have employed SVM ML algorithm for the same.

5.3.3 Chemical Industry

Thermodynamic properties of carbon dioxide are predicted using ML techniques of Back propagation and Regression by Zhien Zhanga et al. [31]. It is of vital importance as CO₂ has a direct impact on green house effect.

5.4 Role of Machine Learning in Training Industry

Machine Learning plays a far impactful role in Training sector both for pedagogy and andragogy schools.

5.4.2 Distance Education

In distance education, it is very vital for the university and teachers to somehow know that amongst all the applicants which one will successfully complete the course till end. An important study done by S. Kotsiantis et al.[32] used ML technique of ensemble classifiers for the same. Their model can predict the performance of students in distance education. This can further be investigated for e-learning or online training mode also.

5.4.2 Student Modeling

Various studies have been done to assess various factors pertaining to students undergoing any training or course. A study on retention capacity of students is done by Dursun Delen[33] using various Machine Learning techniques of SVM, NN and logistic regression methods. A ML based system is proposed by Sentkil Kumar Thangave et al.[34] for placement analysis of students. This model has an accuracy of 71.66% thus giving a scope of improvement and needs further exploration. A ML framework has also been elaborated by A.Pavithra et al.[35] to model students performance.

5.4.3 E-learning vs Traditional methods of training

A survey to compare e learning with traditional method of studying is being carried out by Pena-Ayala et al. [36] using ML methods. In this study they have concluded e learning as being better alternative to traditional one.

6. ANALYSIS AND DISCUSSION

Machine learning paradigms are extensively explored by business community at large but some major challenging areas have also been unearthed during recent ML developments. Some of them are as follows:

6.1 Data Inaccessibility

Large volumes of raw data is required to implement Machine Learning. Not only in training, huge chunks of data are required to model preparation, cleaning and processing. Trust and security issues are involved when dealing with company data. Digital data can be sparse making it difficult to predict data.

6.2 Requirements of infrastructure

Organizations often lack enough infrastructure to facilitate elicitation of data, testing and experimenting with it etc.

6.3 Flexibility of Companies to adapt

Only if a company is agile in its business models it can adapt to the results and changes suggested by a Machine Learning model. The result outcome of a ML paradigm may call for a strategy shift in policies of a company which they may not be ready for.

So the willingness to adapt from the outcome of a ML design approach is crucial in the effectiveness of the ML analysis undertaken.

6.4 Deficit of Talent

There is a lack of machine learning experts. To unlock the true potential of ML methodology right people with expertise in more than one field are required which is generally very sparsely available as of now. According to a survey by Element AI, there are reportedly fewer than 10,000 ML experts in the market who can truly undertake serious research projects.

6.5 Time consuming

Machine learning implementation is way more complicated than usual software development as it involves a lot of data elicitation, training the algorithms with the data, implementing them, and then again learning from the feedback and retraining them according to the suitable rules. Therefore it requires a lot of time and patience for the entire team and the stakeholders.

6.6 Budget Affordability

Running a machine learning project requires a full time data scientist, a project manager, architect, etc all with sound ML knowledge and startups sometimes cannot afford such initial cost. Thus implementing ML in ones organization calls for a revenue analysis and can be a game changer for the organization thinking of employing it.

Computer Scientist Alan Kay said " The best way to predict future is to create it"[37]. Machine learning is going to stay and will continue to develop pushing new horizons in the years to come. [38]. Some open research areas of Machine Learning are summarized for quick reference:

6.7 Agronomics

- Agriculture Robots
- Crop Monitoring and managing
- Soil Monitoring
- Analytic tools for pricing by using social networks of farmers
- data sensors for crop yield and quality

6.8 Ecommerce

- Personalized recommendation engine (eg Netflix, Amazon)
- Personalized content
- Prediction Models for Gaming & Sports
- Prediction model for Dynamic Pricing
- Customer Prediction for purchasing, returning, size and demand etc
- More intensive optimized Search Engines for online business
- Smart chat-bots to improve interaction with customer

6.9 Training & Education

- Developing Smart optimized content using Content Analytics
- Learning Analytics (to track students knowledge and provide content dynamically)
- Intelligent Tutoring
- Virtual Facilitators
- Virtual Learning Environment
- Managing Administration virtually
- Virtual scheduling of trainers with students
- Intelligent grading Systems

6.10 Multi-disciplinary domains of Engineering

- CAD/CAM

- Intelligent Automation
- Image Identification
- Voice Over Identification
- Financial Sector
- Health Sector
- Virtual Assistants
- Prediction Systems
- Videos Monitoring
- Social Media
- Email Hacking
- Filtering Spam
- Online Fraud Detection
- Biometrics
- BioInformatics

As is apparent from the above list, Machine Learning Techniques have a huge potential to steer the future to far greater heights than we can imagine.

CONCLUSION

As the voluminous amount of structured as well as unstructured data is available for analysis, machine learning can play a catalyst to harness this knowledge and convert it into results useful both to the stakeholders and mankind in general. In this paper, a systematic and bibliographic review of the futuristic role of machine learning is presented with emphasis on agronomics, ecommerce, training industry and various other disciplines of manufacturing and development. The need, job and extent of ML as an answer for these new research territories is examined and introduced. Moreover the complex nature of real life problems makes them the best candidate for the Machine Language Solutions by the Researchers, the stakeholders and the end users.

REFERENCES

- [1] Mitchell, T. (1997). Machine Learning. McGraw Hill. p. 2. ISBN 978-0-07-042807-2.
- [2] Samuel AL. Some studies in machine learning using the game of checkers. IBM Journal of research and development. 1959 Jul;3(3):210-29.
- [3] Hayashi, Chikio (1998-01-01). "What is Data Science? Fundamental Concepts and a Heuristic Example". Data Science, Classification, and Related Methods. Studies in Classification, Data Analysis, and Knowledge Organization. Springer Japan. pp. 40–51. doi:10.1007/978-4-431-65950-1_3. ISBN 9784431702085.
- [4] Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.
- [5] Crevier, Daniel (1993), AI: The Tumultuous Search for Artificial Intelligence, New York, NY: BasicBooks, ISBN 0-465-02997-3.
- [6] Samuel, Arthur L. (July 1959), "Some studies in machine learning using the game of checkers", IBM Journal of Research and Development, 3 (3): 210–219.
- [7] McCarthy, John; Minsky, Marvin; Rochester, Nathan; Shannon, Claude (1955), A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, archived from the original on 26 August 2007.
- [8] <http://www.aiia.ed.ac.uk/~dm/dm.html>
- [9] Boyer, R. S., Kaufmann, M., & Moore, J. S. (1995). The Boyer-Moore theorem prover and its interactive enhancement. Computers & Mathematics with Applications, 29(2), 27-62.
- [10] Hopfield, John (April 1982). "Neural networks and physical systems with emergent collective computational abilities" (PDF). Proceedings of the National Academy of Sciences of the United States of America. 79 (8): 2554–2558.
- [11] [11] Marr, Bernard. "A Short History of Machine Learning – Every Manager Should Read". Forbes. Retrieved 28 Sep 2016.
- [12] Harry Henderson (2007). "Chronology". Artificial Intelligence: Mirrors for the Mind. NY: Infobase Publishing. ISBN 978-1-60413-059-1.
- [13] Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. Scientific american, 284(5), 34-43.
- [14] Collobert, Ronan; Benigo, Samy; Mariethoz, Johnny (30 October 2002). "Torch: a modular machine learning software library" (PDF). Retrieved 5 June 2016.
- [15] Moravčík, M., Schmid, M., Burch, N., Lisý, V., Morrill, D., Bard, N., ... & Bowling, M. (2017). Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. Science, 356(6337), 508-513.
- [16] Japkowicz N, Shah M. Performance evaluation in machine learning. In: I El Naqa, R Li, MJ Murphy, eds. Machine Learning in Radiation Oncology: Theory and Applications. Switzerland: Springer-Verlag; 2015: 41– 56.
- [17] Visa, S., Ramsay, B., Ralescu, A. L., & Van Der Knaap, E. (2011). Confusion Matrix-based Feature Selection. MAICS, 710, 120-127.

- [18] Deng, X., Liu, Q., Deng, Y., & Mahadevan, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, 340, 250-261.
- [19] Choubin, Bahram, et al. "River suspended sediment modelling using the CART model: A comparative study of machine learning techniques." *Science of the Total Environment* 615 (2018): 272-281.
- [20] Sajedi-Hosseini, Farzaneh, et al. "A novel machine learning-based approach for the risk assessment of nitrate groundwater contamination." *Science of the Total Environment* 644 (2018): 954-962.
- [21] Shekoofa, Avat, et al. "Determining the most important physiological and agronomic traits contributing to maize grain yield through machine learning algorithms: a new avenue in intelligent agriculture." *PloS one* 9.5 (2014): e97288.
- [22] Pantazi, Xanthoula Eirini, et al. "Wheat yield prediction using machine learning and advanced sensing techniques." *Computers and Electronics in Agriculture* 121 (2016): 57-65.
- [23] Belayneh, A., et al. "Coupling machine learning methods with wavelet transforms and the bootstrap and boosting ensemble approaches for drought prediction." *Atmospheric research* 172 (2016): 37-47.
- [24] El-Bendary, Nashwa, et al. "Using machine learning techniques for evaluating tomato ripeness." *Expert Systems with Applications* 42.4 (2015): 1892-1905.
- [25] Mori, Junichiro, et al. "Machine learning approach for finding business partners and building reciprocal relationships." *Expert Systems with Applications* 39.12 (2012): 10402- 10407.
- [26] Bigus, J. P. (1996). *Data mining with neural networks: solving business problems from application development to decision support*, McGraw-Hill, Inc.
- [27] He, Yuqing, Kamaladdin Fataliyev, and Lipo Wang. "Feature selection for stock market analysis." *International Conference on Neural Information Processing*. Springer, Berlin, Heidelberg, 2013.
- [28] Oliveira, Adriano LI, et al. "GA-based method for feature selection and parameters optimization for machine learning regression applied to software effort estimation." *information and Software Technology* 52.11 (2010): 1155-1166.
- [29] Malhotra, Ruchika, and Ankita Jain. "Fault prediction using statistical and machine learning methods for improving software quality." *Journal of Information Processing Systems* 8.2 (2012): 241-262.
- [30] Senanayaka, Jagath Sri Lal, et al. "Early detection and classification of bearing faults using support vector machine algorithm." *2017 IEEE Workshop on Electrical Machines Design, Control and Diagnosis (WEMDCD)*. IEEE, 2017.
- [31] Zhang, Zhien, et al. "Machine learning predictive framework for CO₂ thermodynamic properties in solution." *Journal of CO₂ Utilization* 26 (2018): 152-159.
- [32] Kotsiantis, S., Kiriakos Patriarcheas, and M. Xenos. "A combinational incremental ensemble of classifiers as a technique for predicting students' performance in distance education." *Knowledge-Based Systems* 23.6 (2010): 529-535.
- [33] Delen, Dursun. "A comparative analysis of machine learning techniques for student retention management." *Decision Support Systems* 49.4 (2010): 498-506.
- [34] Thangavel, Sentkil Kumar, P. Divya Bkaratki, and Abijitk Sankar. "Student placement analyzer: A recommendation system using machine learning." *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*. IEEE, 2017.
- [35] A.Pavithra , S.Dhanaraj. "Prediction Accuracy on Academic Performance of Students Using Different Data Mining Algorithms with Influencing Factors." *International Journal of Scientific Research in Computer Science Applications and Management Studies*. 7.5 (2018).; 1-7
- [36] Peña-Ayala, Alejandro. "Educational data mining: A survey and a data mining-based analysis of recent works." *Expert systems with applications* 41.4 (2014): 1432-1462.
- [37] https://en.wikiquote.org/wiki/Alan_Kay
- [38] Zhu, G., Liu, D., Du, Y., You, C., Zhang, J., & Huang, K. (2020). Toward an intelligent edge: wireless communication meets machine learning. *IEEE Communications Magazine*, 58(1), 19-25.