A Review on Machine Learning Algorithm and Artificial Intelligence

Jai Santani[1]

[1] B.tech in Computer Science Engineering Department, Technocrats Institute of Technology, Bhopal, Madhya Pradesh.

Abstract— A long time ago, there were various tasks that could only be done by humans. There is no machine and technology like today. At that time, science had not yet developed and technology had not been invented. Therefore, work depends entirely on people and humanity has realized that "today 's science is the technology of tomorrow". Advanced new technology is no less than God's blessing. Adaptive inventions have been invented that reduce human jobs and bright prospects, referred to as artificial intelligence and machine learning. Although there were many false assumptions in the beginning, we are witnessing a new era of error-free technology and excellent science. This critique deals with general concepts of artificial intelligence and machine learning.

Index Terms—:   reasoning, general AI, supervised learning, Machine learning, algorithms, pseudo code, artificial intelligence.

1.   INTRODUCTION

(I) INTRODUCTION: Just as humans have proven their existence on earth, it is important for everyone to understand what artificial intelligence and machine learning mean to humans. This is a good poem summarized by WH Auden, which gives the relationship between human life and people. robot technology:

Those who will not speak

It will be lost in practice.

Those who will not act

It will be lost for this reason

The poem explains that the "Fittest survival", ie only those people / machines will survive that prove their existence with the best performance, high intelligence and maximum capacity. Hence its high weather and one of the most wonderful discoveries of excellence, since man learned to create tools and fire. It is a road that we just walk forward. there will be no delay, as soon as we achieve the super-intelligence of the machine which will be self-learning, fully automatic and self-improved.

As we have seen in science fiction movies like Iron Man, Star Wars, Terminator, etc. Destroy all the super robots around the world in terms of fully equipped superheroes. Even if someone has seen rockets or space shuttle launches, motor cars or just robots can help you complete household chores and complete basic full muscle exercises. This is a good example of human intelligence. There is no doubt that artificial intelligence and machine learning are the two most popular keywords in the world and are often used interchangeably.

(II) MACHINE LEARNING : INTERSECTION OF STATISTICS AND COMPUTER SCIENCE:

Machine Learning was the apparent result when Computer Science and Statistics came together. Computer Science focuses on building machines that solve specific problems and tries to identify if problems are solvable. The basic approach used by Statistics is basically data mining, case modeling and measuring the reliability of conclusions. The defining idea of Machine Learning is a little different but in part depends on both. While Computer Science focuses on manual computer programming, MLad faces the problem of causing computers to reprogram themselves each time they are exposed to new data based on some of the original learning strategies provided. Statistics, on the other hand, focus on data mining and probability, and Machine Learning includes additional concerns about the feasibility and effectiveness of architectures and algorithms for processing this data, combining multiple learning tasks into a compact and performance measure.

(III) MACHINE LEARNING AND HUMAN LEARNING

A third area of research closely related to Machine Learning is the study of the human and animal brains in Neuroscience, Psychology and related fields. The researchers suggested that how a machine could learn from experience would probably not be significantly different from how an animal or human mind learns with time and experience. However, research focused on solving machine learning problems using human brain learning methods has not yielded very promising results so far from research related to the statistical-computational approach. This may be due to the fact that human or animal psychology remains to this day fully understood. Regardless of these difficulties, the collaboration between human learning and machine learning is increasing for machine learning is used to explain various learning techniques seen in humans or animals. For example, a time difference machine learning method has been proposed to explain neural signals in animal learning. It is quite expected that this cooperation will increase significantly in the coming years.

(IV) DATA MINING, ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

In practice, these three branches are so interconnected and overlapping that they almost draw a boundary or hierarchy between the three. In other words, these three areas are symbiotically related and a combination of these approaches can be used as a tactic to produce more effective and sensitive results. Approximately, data mining is essentially about interpreting any kind of data, but it lays the foundation for both artificial intelligence and machine learning. In practice, it is not just a sample of information from various sources, but it analyzes and recognizes patterns and correlations that exist in this information that would be difficult to interpret manually. Therefore, data mining is not a simple method of proving a hypothesis, but a method of making relevant hypotheses. This extracted data and the corresponding patterns and assumptions can be used as a basis for both machine learning and artificial

intelligence. Artificial intelligence can be broadly defined as a machine that has the ability to solve a given problem on its own without any human intervention. Solutions are not programmed directly into the system, but the necessary data and the AI that interprets this data produce a solution on its own. The following interpretation is just a data mining algorithm.

II. TRADITIONAL METHODS

The main contributions of the authors are twofold. First, we provide a new set of benchmarks for the predicted accuracy of machine learning methods in measuring total market risk premiums and individual stocks. This accuracy can be summarized in two ways. The first is a prediction of a high off-sample R2 compared to previous literature that is strong in a variety of machine learning specifications. Second, and most importantly, we present great financial benefits to investors who use machine learning forecasts. A portfolio strategy that multiplies the S&P 500 with neural network forecasts enjoys an annual Sharpe index of 0.77 versus a Sharpe 0.51 ratio of a buy-and-hold investor. And a long-term decimal spread strategy that takes positions based on neural network forecasts at the stock level earns an annual Sharpe ratio outside the 1.35 sample, more than doubling the performance of a bottom-down regression strategy. [1]

Today's telecommunications networks have become sources of vast heterogeneous data. This information can be retrieved from network traces, network alarms, signal quality indicators, user behavior data, etc. Advanced mathematical tools are required to extract important information from this data and to make decisions about the proper functioning of the networks from the data generation network. Among these mathematical tools, Machine Learning (ML) is considered one of the most promising methodological analytics for network data analysis and allows for automated network self-configuration and error management. [2]

In this article, we will describe tasks and applications for machine learning, as well as the potential threats associated with current data collection methods or the construction of ML systems. We will further elaborate on the techniques proposed for protecting the privacy of individuals or companies. Our intention is to bridge the gap between machine learning and privacy and security technologies, helping professionals in both sectors become more familiar with machine learning, the potential threats to privacy, the proposed solutions and challenges that will follow. [3]

Diabetes is one of the leading causes of blindness, kidney failure, amputation, heart failure and stroke. When we eat, our body converts food into sugars or glucose. At this point, our pancreas is supposed to release insulin. Insulin serves as the "key" to opening our cells, allowing glucose to enter - and allowing us to glucose for energy. But with diabetes, this system does not work. Many important things can go wrong - causing diabetes. Type 1 and 2 diabetes are the most common forms of the disease, but there are other types, such as gestational diabetes, which occurs during pregnancy, as well as other forms. This work focuses on recent developments in machine learning that have significant implications for the detection and diagnosis of diabetes. [4]

Deep reinforcement learning is ready to revolutionize the field of artificial intelligence and represents a step towards building autonomous systems with a higher level of understanding of the visual world. Currently, in-depth learning allows you to learn scaling-up on problems that were previously inappropriate, such as learning to play video games directly from pixels. Deep gain learning algorithms also apply to robotics, allowing robot control policies to be learned directly from real-world camera inputs. In this research, we begin with an introduction to the general field of reinforcement learning and then move on to the mainstream value-based and policy-based methods. Our research will cover key algorithms in deep reinforcement learning, such as the deep Q network, trust area policy optimization, and the asynchronous actor-critic advantage. At the same time, we highlight the unique benefits of deep neural networks, focusing on visual understanding through reinforcement learning. In conclusion, we describe several current areas of research in the field. [5]

In recent decades, Machine Learning (ML) has evolved from the attempt of a few computer enthusiasts to take advantage of the ability of computers to learn to play games and a part of Mathematics (Statistics) that rarely looked at computational approaches, to an independent research discipline that has not only provided the necessary basis for the statistical-computational principles of learning processes, but has also developed various algorithms that are regularly used for text interpretation, pattern recognition and many other commercial purposes and has led to a particular research interest in locating data mining. hidden irregularities or irregularities in social data that increase with the latter. This article focuses on explaining the concept and evolution of Machine Learning, some of the most popular Machine Learning algorithms, and attempts to compare three most popular algorithms based on some key concepts. The Sentiment140 dataset was used and the performance of each algorithm in terms of training time, prediction time and prediction accuracy have been documented and compared. [6]

Many days ago, there was all kinds of work done only by humans. There were no such machines and technologies as today. At that time, science was not developed and technologies were not invented. So work is completely dependent on the people and people have recognized that "today's science is tomorrow's technology". New advanced technologies are no less than a blessing from God Adaptive inventions were invented to reduce human work and a bright future called simply Artificial Intelligence and Machine Learning. Although there were many false assumptions in the beginning, we are witnessing a new era without technology and superior science. This review covers the general concepts of artificial intelligence and machine learning. [7]

Supervised machine learning algorithms are those algorithms that need external assistance. The input data set is divided into train and test data sets. The train data set has an output variable that must be predicted or sorted. All algorithms learn some kind of template from the training dataset and apply it to the test dataset for prediction or classification. [8]

The target audience for this seminar is those who want to get started quickly in the field of data science and machine learning. We will provide an overview of current and most popular libraries with an emphasis on Python, however we will list alternatives in other languages where needed. All tools presented here are free and open source, and many are licensed under very flexible terms (including, for example, commercial use). Each library will be inserted, a code will be displayed and typical usage cases will be described. Medical datasets will be used to demonstrate many of the algorithms. [9]

Machine learning (ML) and technical intelligence (AI) are becoming dominant problem-solving techniques in many areas of research and industry, mainly due to the recent success of deep learning (DL). However, the equation AI = ML = DL, as recently proposed in news, blogs and media, is very small. These fields share the same fundamental assumptions: computation is a useful way of modeling intelligent behavior on machines. What kind of calculation and how to program it? This is not the right question. The calculation does not exclude search methods, logical and probabilistic, nor learning methods (deep) (un) under supervision and reinforcement, among others, as the computational models include all of them. They complement each other, and the next important discovery lies not only in the impulse of each but also in their combination. [10]

In this article, the authors combine text mining and machine learning techniques to support the identification of the most relevant literature using a collection of datasets searched in many journal repositories: ACM Portal 1, Engineering Village 2, IEEE Xplore 3, ScienceDirect 4, Web of Science 5. The data set is quantified in order to reduce the time used to review the literature review papers for our research area: high-performance computing in support of computer-aided diagnostic systems using medical images. Text mining is a common process of extracting relevant information using a set of documents. It provides basic preprocessing methods, such as authentication, extraction of representative features, and advanced functions as a set of complex patterns. Document sorting is the task of assigning text to one or more categories: the object class name and the main topics. [11]

This paper provides an extensive review of studies related to expert evaluation for software development using Machine Learning Techniques (MLT). Machine learning in this new age proves the promise of consistently producing accurate estimates. The machine learning system "learns" how to be evaluated by all the completed projects. The main purpose and contribution of the review is to support research for expert evaluation, ie to facilitate other researchers for relevant expert evaluation studies using machine learning techniques. This article introduces the most commonly used machine learning techniques, such as neural networks, hypothesis reasoning, classification and regression trees, rule induction, genetic algorithm, and genetic programming for expert software development. In each of our studies we found that the results of different machine learning techniques depend on the areas of application in which they are applied. Our review of the study not only shows that these techniques are competitive with traditional appraisers in a data set, but also shows that these methods are sensitive to the data in which they are trained. [12]

Much of current machine learning (ML) research has lost its connection to problems of introduction to the wider world of science and society. From this perspective, there are obvious limitations to the data sets we explore, the metrics we use for evaluation, and the extent to which the results are reported back to their original domains. What changes are needed in the way research is conducted to increase the impact that ML has? We present six Impact challenges to explicitly focus the energy and attention of the sector and discuss the existing obstacles that need to be addressed. Our goal is to inspire ongoing discussion and focus on the ML that matters. [13]

With the increasing availability of electronic documents and the rapid development of the World Wide Web, the task of automatically categorizing documents has become the basic method for organizing the discovery of information and knowledge. Proper classification of electronic documents, online news, blogs, emails and digital libraries requires text mining, machine learning and natural language editing techniques to gain essential knowledge. The aim of this paper is to highlight the important techniques and methodologies used in text document classification, while at the same time realizing some of the interesting challenges that need to be solved, focusing mainly on text representation and machine learning techniques. This paper provides an overview of the theory and methods of document classification and text mining, focusing on existing literature. [14]

Supervised machine learning is the search for algorithms that justify externally provided instances to produce general assumptions, which then make predictions for future cases. In other words, the goal of supervised learning is to build a concise class tag distribution model in terms of predictive characteristics. The resulting classifier is then used to assign class tags in test cases where the values of the prediction attributes are known but the value of the class tag is unknown. This paper describes various supervised classification techniques for machine learning. Of course, a single article may not be a complete overview of all supervised machine learning classification algorithms (also known as induction classification algorithms), but we hope that the references cited will address the most important theoretical issues, guiding the researcher to interesting and interesting research combinations of prejudice that have not yet been explored. [15]

This work introduces FINKNN, a k-nearest neighbor classifier that operates over the metric grid of conventional space-supported convex fuzzy sets. We show that for problems concerning measurement populations, the data can be represented by fuzzy interval numbers (FIN) and we present an algorithm for constructing FIN from such populations. We then present a small metric measurement between FINs with arbitrary membership functions, which forms the basis for FINKNN similarity measurements. We apply FINKNN in the project of forecasting the annual sugar production based on population measurements provided by the Hellenic Sugar Industry. We show that FINKNN improves the accuracy of predictions in this work and discuss the broader scope and potential usefulness of these techniques. [16]
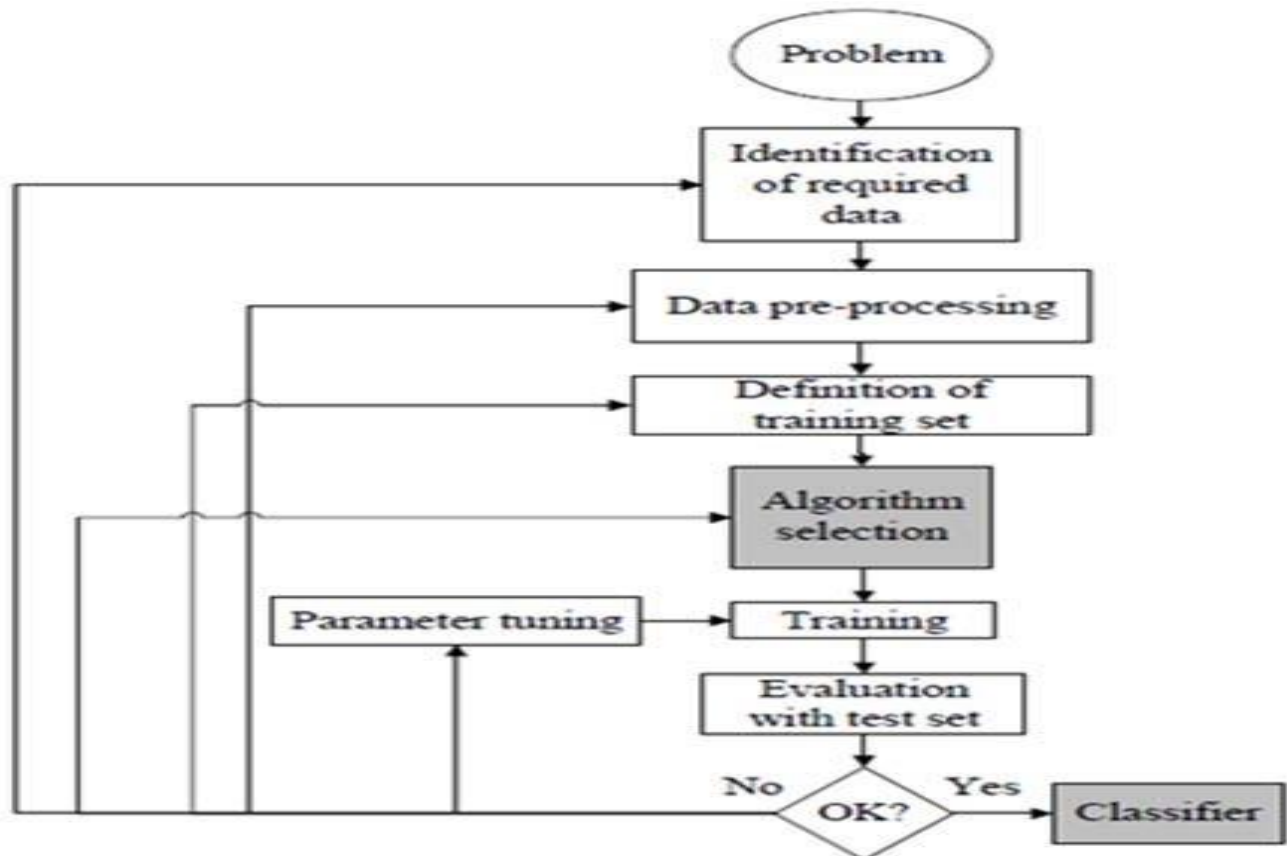
III. METHODOLOGY

Figure 1: Workflow of supervised machine learning algorithm

Supervised machine learning algorithms are those algorithms that need external assistance. The input data set is divided into train and test data sets. The train data set has an output variable that must be predicted or sorted. All algorithms learn some kind of template from the training data set and apply it to the test data set for prediction or classification. The workflow of the supervised machine learning algorithms is given in Fig. 1.
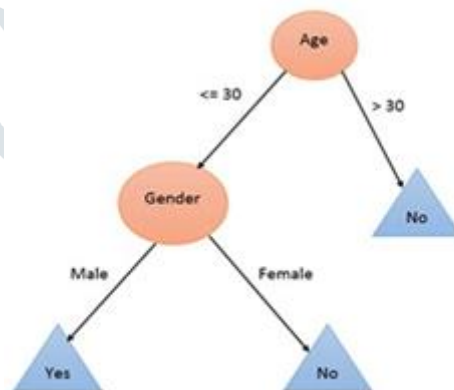

Figure 2: Decision Tree

Decision trees are these types of trees that group features by classifying them based on their prices. The decision tree is mainly used for classification purposes. Each tree consists of nodes and branches. Each node typically represents a group to be classified, and each branch represents a value that the node can take. An example decision tree is given in Figure 2.

The pseudo code for the decision tree is described in Fig. 4. Where S, A and y are training set, input attribute and target attribute respectively.
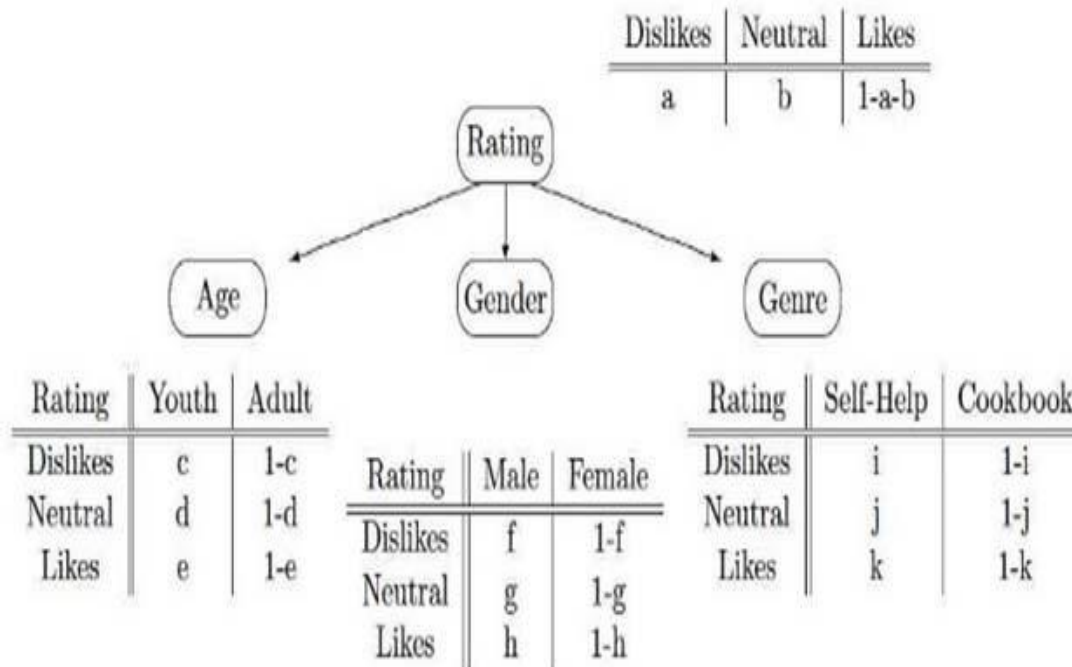
Dislikes	Neutral	Likes
a	b	1-a-b

Rating

Age   Gender   Genre

Rating	Youth	Adult
Dislikes	c	1-c
Neutral	d	1-d
Likes	e	1-e

Rating	Male	Female
Dislikes	f	1-f
Neutral	g	1-g
Likes	h	1-h

Rating	Self-Help	Cookbook
Dislikes	i	1-i
Neutral	j	1-j
Likes	k	1-k

Figure 3: An Example of Bayesian Network

Naïve Bayes mainly targets the text sorting industry. It is mainly used for grouping and classification purposes. The underlying architecture of Naïve Bayes depends on conditional probability. Creates trees based on the probability that they will occur. These trees are also known as the Bayesian Network. An example of the network is given in Fig. 3. The pseudo code is given in Figure 4.

```
procedure DTInducer(S, A, y)
1:  T = TreeGrowing(S, A, y)
2:  Return TreePruning(S,T)
procedure TreeGrowing(S, A, y)
1:  Create a tree T
2:  if One of the Stopping Criteria is fulfilled then
3:      Mark the root node in T as a leaf with the most common
        value of y in S as the class.
4:  else
5:      Find a discrete function f(A) of the input attributes val-
        ues such that splitting S according to f(A)'s outcomes
        (v₁,...,vₙ) gains the best splitting metric.
6:      if best splitting metric ≥ treshold then
7:          Label the root node in T as f(A)
8:          for each outcome vᵢ of f(A) do
9:              Subtreeᵢ = TreeGrowing(σ_{f(A)=vᵢ}S, A, y).
10:             Connect the root node of T to Subtreeᵢ with an
                edge that is labelled as vᵢ
11:         end for
12:     else
13:         Mark the root node in T as a leaf with the most
            common value of y in S as the class.
14:     end if
15: end if
16: Return T
procedure TreePruning(S, T, y)
1:  repeat
2:      Select a node t in T such that pruning it maximally
        improve some evaluation criteria
3:      if t ≠ ∅ then
4:          T = pruned(T, t)
5:      end if
6:  until t=∅
7:  Return T
```

Figure 4: Pseudo code for Naïve Bayes

IV. Conclusion and Future Scope

Conclusion

The whole world is on the path of Digitization and for this purpose the concept of artificial intelligence and machine learning plays an important role. Our research paper is based entirely on how intelligence and new machine technologies are invented in our daily lives. Today's machines are ready to train knowledge and are responsible for improving intelligence. In the future, we do not think and imagine the progress of the world due only to Artificial Intelligence and Innovative Machines. We cannot imagine what is happening in the environment and around the world due to scientists and engineers. The scientist has developed robots that work like human beings and also the research continues to create the best world in the future. Supporting the new generation is one of the most important parts of developing new technologies. The combination of Science and Engineering and quality machine learning will surely lead the world to the highest possible adaptation.

This article explores various machine learning algorithms. Today, every person uses machine learning knowingly or unknowingly. From receiving a suggested product in online shopping to updating photos on social networking sites. This article gives an introduction to most of the popular machine learning algorithms.

AUTHORS REVIEW

Author review many papers and he found Design & Development of a Hybrid Algorithm of machine learning using workflow of supervised machine learning algorithm, decision tree, Bayesian network and pseudo code.

REFERENCES

[1.] Shihao Gu, Bryan Kelly, Bryan Kelly, Dacheng Xiu "Empirical Asset Pricing via Machine Learning" SFS 2020.

[2.] Francesco Musumeci, Cristina Rottondi, Avishek Nag, Irene Macaluso, Darko Zibar, Marco Ruffini, Massimo Tornatore "An Overview on Application of Machine Learning Techniques in Optical Networks" IEEE 2018.

[3.] Mohammad Al-Rubaie, J. Morris Chang "Privacy Preserving Machine Learning: Threats and Solutions" IEEE 2018.

[4.] Priyanka Indori, Yogesh Kumar Rathore "A Survey: Detection and Prediction of Diabetes Using Machine Learning Techniques" IJERT 2018.

[5.] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, Anil Anthony Bharath "A Brief Survey of Deep Reinforcement Learning" IEEE 2017.

[6.] Kajaree Das, Rabi Narayan Behera "A Survey on Machine Learning: Concept, Algorithms and Applications" International Journal of Innovative Research in Computer and Communication Engineering 2017.

[7.] Mahind Rupali, Patil Amit "A Review Paper on General Concepts of Artificial Intelligence and Machine Learning" IARJSET 2017.

[8.] Ayon Dey "Machine Learning Algorithms" IJCSIT 2016.

[9.] Marcus D. Bloice and Andreas Holzinger "A Tutorial on Machine Learning and Data Science Tools with Python" Research Gate 2016.

[10.]   Kristian Kersting "Machine Learning and artificial intelligence" SPECIALTY GRAND CHALLENGE 2016.

[11.]   CarlosA.S.J. Gulo1, ThiagoR.P.M. Rúbio, Shazia Tabassum, and SimoneG.D. Prado "Mining Scientific Articles Powered by Machine Learning Techniques" iccsw15 2015.

[12.]   Yogesh Singh, Pradeep Kumar Bhatia, Omprakash Sangwan "A REVIEW OF STUDIES ON MACHINE LEARNING TECHNIQUES" International Journal of Computer Science and Security 2013.

[13.]   Kiri L. Wagstaff "Machine Learning that Matters" International Conference on Machine Learning, Edinburgh, Scotland 2012.

[14.]   Aurangzeb Khan, Baharum Baharudin, Lam Hong Lee*, Khairullah khan "A Review of Machine Learning Algorithms for Text-Documents Classification" JOURNAL OF ADVANCES IN INFORMATION TECHNOLOGY 2010.

[15.]   S. B. Kotsiantis "Supervised Machine Learning: A Review of Classification Techniques" Informatica 2007.

[16.]   Vassilios Petridis, Vassilis G. Kaburlasos "FINkNN: A Fuzzy Interval Number k-Nearest Neighbor Classifier for Prediction of Sugar Production from Populations of Samples" Journal of Machine Learning Research 2003.