Machine Learning Algorithms and Applications: A Review

Gaurav Mandloi
Assistant Professor
Department of Information Technology
Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India

Abstract— In past few decades, Machine Learning (ML) has evolved from the endeavor of few computer enthusiasts exploiting the possibility of computers learning to play games, and a part of Mathematics (Statistics) that seldom considered computational approaches, to an independent research discipline that has not only provided the necessary base for statistical-computational principles of learning procedures, but also has developed various algorithms which are regularly used for text interpretation, pattern recognition, and a many other commercial purposes and has led to a separate research interest in mining and machine learning to find hidden regularities or irregularities in social data that growing by second. This paper focuses on study of various algorithm and applications of Machine Learning, which are used for various purposes like data mining, image processing, predictive analytics, etc.

Keywords: Machine Learning, Pattern Recognition, Algorithm, Data, Training, accuracy

I. INTRODUCTION

We know that 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 and designing of algorithms that can learn from and make predictions on data sets. These methods operate by construction of a model from example inputs in order to make data-driven predictions or choices rather than following firm static program instructions.

"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." -- Tom Mitchell, Carnegie Mellon University.

So if we want our program to foresee, for example, traffic forms at a busy node (task T), we can run it through a machine learning process with data about previous traffic patterns (experience E) and, if it has successfully "learned", it will then do better at predicting upcoming traffic patterns (performance measure P). We require machine learning in the following cases:

- Human expertise is absent. E. g. Navigating on Mars.
- Humans are unable to explain their expertise. E. g. Speech Recognition.
- Solution changes with time E.g. Temperature Control.
- Solution needs to be adapted to particular cases. E. g. Biometrics.
- Problem size is too vast for our limited reasoning capabilities. E. g. Calculating webpage ranks.

Machine leaning involves two types of tasks:

1) Supervised Machine Learning: The program is "trained" on a pre-defined set of "training examples", which then facilitate its ability to reach an accurate conclusion when given new data.
2) Unsupervised Machine Learning: The program is given a bunch of data and must find patterns and relationships therein.

II. TYPES OF LEARNING

A. Supervised Learning

The supervised machine learning algorithms are those algorithms which needs external assistance. The input dataset is divided into training and test dataset. The training dataset has output variable which needs to be predicted or classified. All algorithms learn some kind of patterns from the training dataset and apply them to the test dataset for prediction or classification [4]. The workflow of supervised machine learning algorithms is given in Fig. 2. Three most famous supervised machine learning algorithms have been discussed here.

1) Decision Tree:

Decision trees are those type of trees which groups attributes by sorting them based on their values. Decision tree is used mainly for classification purpose. Each tree consists of nodes and branches. Each nodes represents attributes in a group that is to be classified and each branch represents a value that the node can take [4]. Example of decision tree is given in Fig. 2.

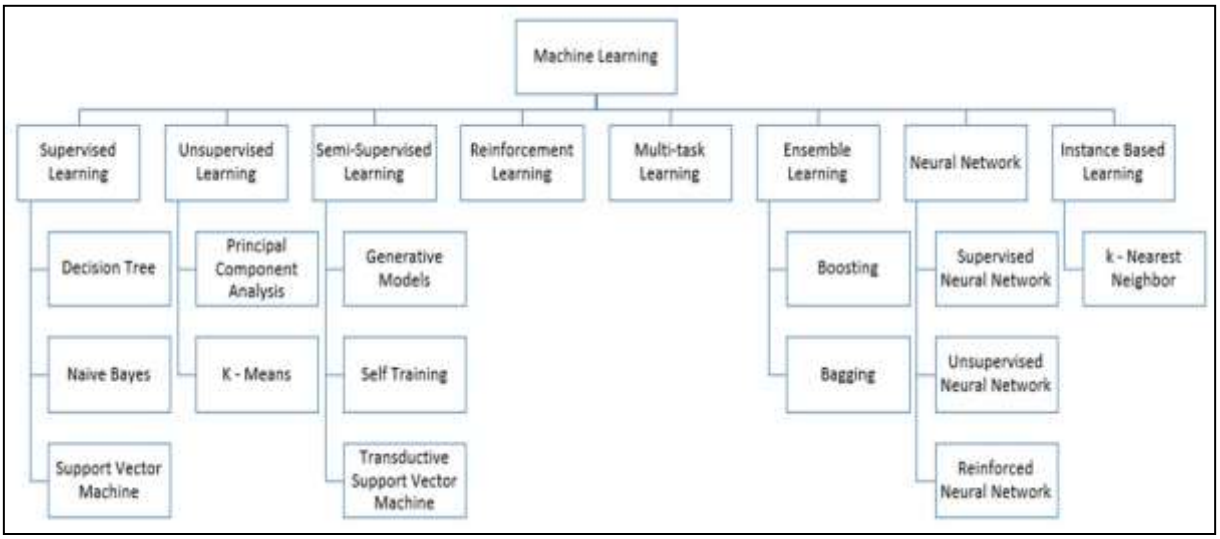

Fig. 1: Types of Learning [2] [3]

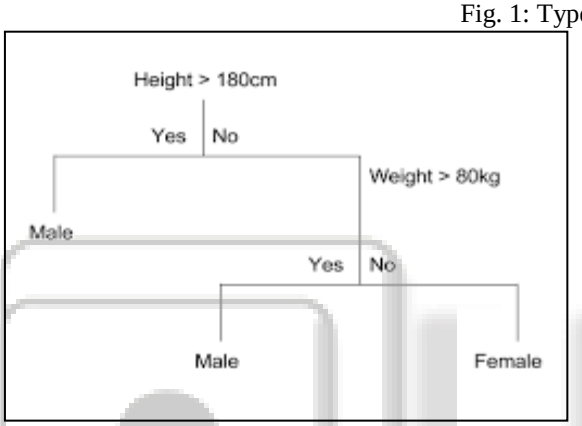

Fig. 2: Decision Tree

2) Naïve Bayes:

Naïve Bayes mainly targets the text classification industry. This Byes is mainly used for clustering and classification purpose [6]. The underlying architecture of Naïve Bayes depends on the conditional probability. It generates trees based on their probability of happening. These trees are also known as Bayesian Network. An example of the network is given in Fig. 3.

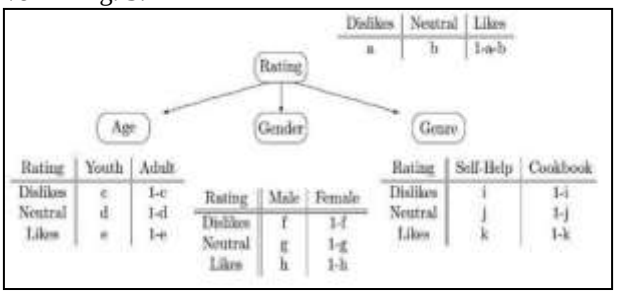

Fig. 3: An Example of Bayesian Network [7]

3) Support Vector Machine:

Another most widely used state-of-the-art machine learning technique is Support Vector Machine (SVM). It is mainly used for classification. Support Vector Machine works on the principle of margin calculation. It basically, draw margins between the classes. The margins are drawn in such a fashion that the distance between the margin and the classes is maximum and hence, minimizing the classification error. An example of working of SVM is given in Fig. 4.

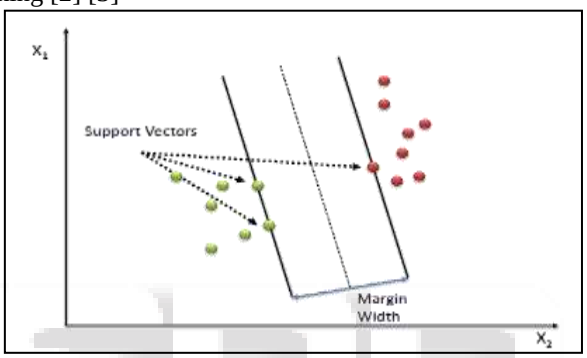

Fig. 4: Working of Support Vector Machine [8]

B. Unsupervised Learning

The unsupervised learning algorithms learns few features from the data. When new dataset is introduced, it uses the previously learned features to recognize the class of the data. This learning is mainly used for clustering and feature reduction. Workflow of unsupervised learning is given in Fig. 5.

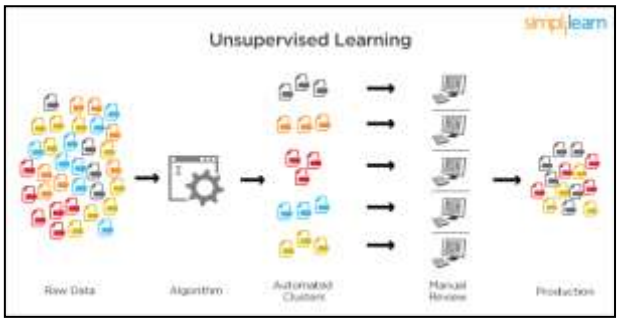

Fig. 5: Example of Unsupervised Learning [10]

The two main algorithms for clustering and dimensionality reduction techniques are discussed below.

1) K-Means Clustering:

Clustering or grouping is a type of unsupervised learning technique that when initiates, creates groups automatically. The items which possesses similar characteristics are put in the same cluster. The algorithm is called k-means because it creates k distinct clusters. The mean values in a particular cluster is the center of that cluster [9]. A clustered data is represented in Fig. 6.

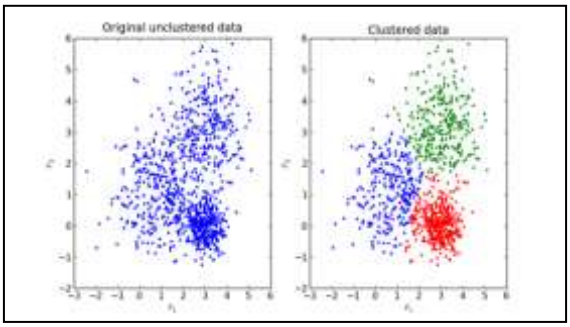

Fig. 6: K-Means Clustering [12]

2) Principal Component Analysis:

In Principal Component Analysis or PCA, the dimension of the data is reduced to make the computations faster and easier. To understand how PCA works, let's take an example of 2D data. When data is being plot in a graph, it will take up two axes. PCA is applied on the data, the data then will be 1D. This is explained in Fig. 7.

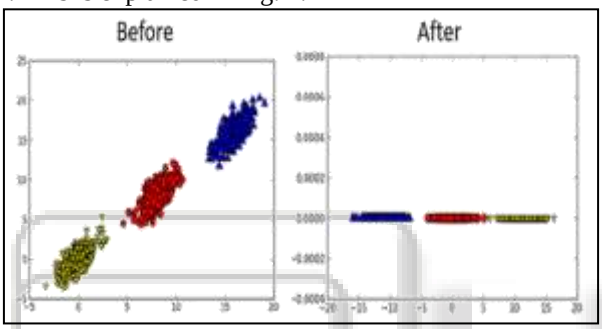

Fig. 7: Visualization of data before and after applying PCA [11]

C. Semi - Supervised Learning

Semi – supervised learning algorithms is a technique which combines the power of both supervised and unsupervised learning. It can be fruit-full in those areas of machine learning and data mining where the unlabeled data is already present and getting the labeled data is a tedious process [15]. There are many categories of semi-supervised learning [16]. Some of which are discussed below:

1) Generative Models:

Generative models are one of the oldest semi-supervised learning method assumes a structure like $p(x,y) = p(y)p(x|y)$ where $p(x|y)$ is a mixed distribution e.g. Gaussian mixture models. Within the unlabeled data, the mixed components can be identifiable. One labeled example per component is enough to confirm the mixture distribution.

2) Self-Training:

In self-training, a classifier is trained with a portion of labeled data. The classifier is then fed with unlabeled data. The unlabeled points and the predicted labels are added together in the training set. This procedure is then repeated further. Since the classifier is learning itself, hence the name self-training.

3) Transductive SVM:

Transductive support vector machine or TSVM is an extension of SVM. In TSVM, the labeled and unlabeled data both are considered. It is used to label the unlabeled data in such a way that the margin is maximum between the labeled and unlabeled data. To find an exact solution by TSVM is a NP-hard problem.

D. Reinforcement Learning

Reinforcement learning is a type of learning which makes decisions based on which actions to take such that the outcome is more positive. The learner has no knowledge which actions to take until it's been given a situation. The action which is taken by the learner may affect situations and their actions in the future. Reinforcement learning completely depends on two criteria: trial and error search and delayed outcome [17]. The general model [18] for reinforcement learning is depicted in Fig. 8.

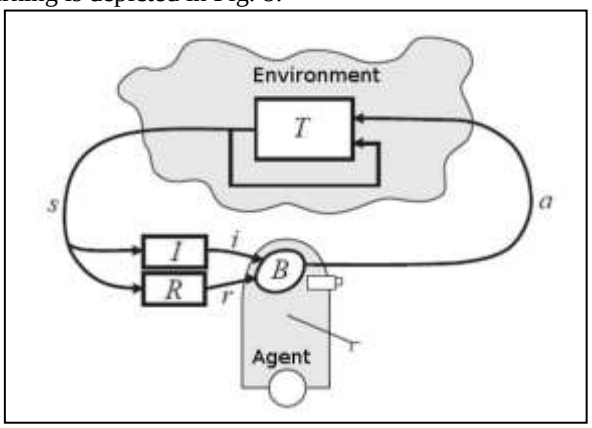

Fig. 8: The Reinforcement Learning Model [18]

In the figure, the agent receives an input i, current states, state transition r and input function I from the environment. Based on these inputs, the agent generates a behavior B and takes an action a generates an outcome.

E. Multitask Learning

Multitask learning has a simple goal of helping other learners to perform better. When multitask learning algorithms are applied on a task, it remembers the procedure how it solved the problem or how it reaches to the particular conclusion. Algorithm then uses these steps to find the solution of other similar problem or task. This type of helping of one algorithm to another can also be termed as inductive transfer mechanism. If the learners share their experience with each other, the learners can learn concurrently rather than individually and can be much faster [19].

F. Ensemble Learning

When various individual learners are combined to form only one learner then that particular type of learning is called ensemble learning. The individual learner may be Naïve Bayes, decision tree, neural network, etc. Ensemble learning is a hot topic since 1990s. It has been observed that, a collection of learners is almost always better at doing a particular job rather than individual learners [20]:

1) Boosting:

Boosting is a technique in ensemble learning which is used to decrease bias and variance. Boosting generates a collection of weak learners and convert them to one strong learner. Weak learner is a classifier that is barely correlated with true classification. On the other side, a strong learner is a type of classifier which is strongly correlated with true classification [21].

2) Bagging:

Bagging or bootstrap aggregating is applied where the accuracy and stability of a machine learning algorithm needs

to be increased. It is applicable in classification and regression. Bagging also decreases variance and helps in handling overfitting [23].

G. Neural Network Learning

Artificial Neural network (ANN) is derived from the biological concept of neurons. A neuron is a cell like structure in a brain. To understand neural network, one must understand how a neuron works. A neuron has mainly four parts (see Fig. 9). They are dendrites, nucleus, soma and axon.

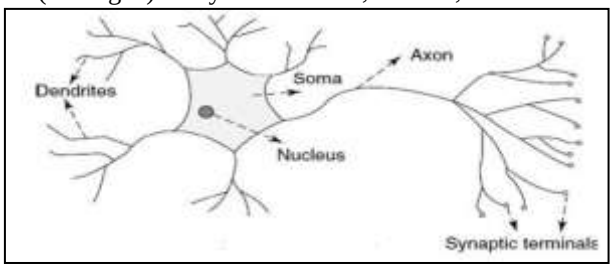

Fig. 9: A Neuron [24]

The dendrites receive electrical signals. Soma processes the electrical signal. The output of the process is carried by the axon to the dendrite terminals where the output is sent to next neuron. The nucleus is the heart of the neuron. The inter connection of neuron is called neural network where electrical impulses travel around the brain.

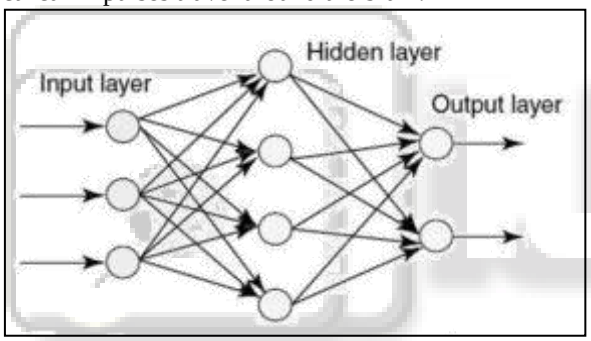

Fig. 10: Structure of an Artificial Neural Network [24]

An artificial neural network behaves the same way. It works on three layers. The first layer i.e. input layer takes input (much like dendrites). The middle laye i. e. hidden layer processes the input (like soma and axon). Finally, the output layer sends the calculated output (like dendrite terminals) [24]. There are basically three types of artificial neural network: supervised, unsupervised and reinforcement [25].

1) Supervised Neural Network:

In the supervised neural network, the output of the input is already known. The predicted output of the neural network is compared with the actual output. On the basis of the error, the parameters are changed, and then fed into the neural network again. Fig. 11 will summarize the process. Supervised neural network is used in feed forward neural network.

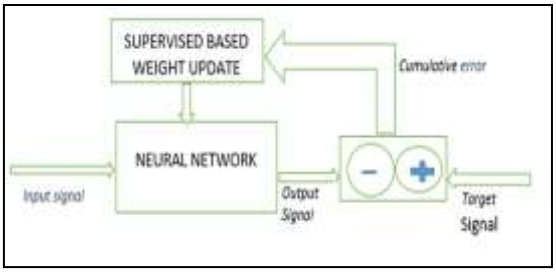

Fig. 11: Supervised Neural Network [25]

2) Unsupervised Neural Network:

Here, the neural network has no prior clue about the output the input. The main job of the network is to categorize the data according to some similarities. The neural network checks the correlation between various inputs and groups them. The schematic diagram is shown in Fig. 12.

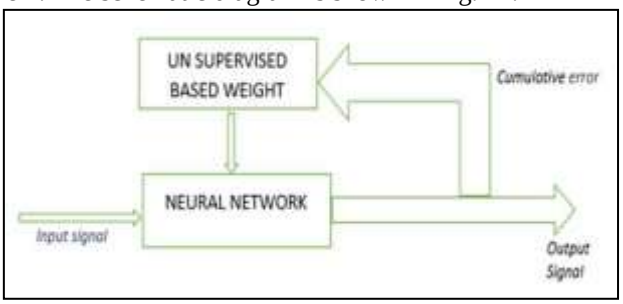

Fig. 12: Unsupervised Neural Network [25]

3) Reinforced Neural Network:

In reinforced neural network, the network behaves as if a human communicates with the environment. From the environment, a feedback has been provided to the network acknowledging the fact that whether the decision taken by the network is right or wrong. If the decision is right, the connections which points to that particular output is strengthened. The connections are weakened otherwise. The network has no previous information about the output. Reinforced neural network is represented in Fig. 13.

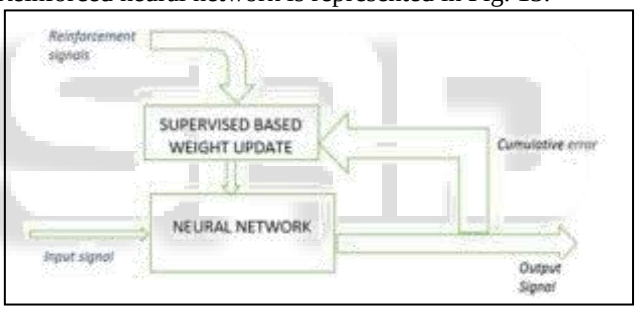

Fig. 13: Reinforced Neural Network [25]

H. Instance-Based Learning

In instance-based learning, the learner learns a particular type of pattern. It tries to apply the same pattern to the newly fed data. Hence the name instance-based. It is a type of lazy learner which waits for the test data to arrive and then act on it together with training data. The complexity of the learning algorithm increases with the size of the data. Given below is a well-known example of instance-based learning which k-nearest neighbor [26] is.

1) K-Nearest Neighbor:

In k-nearest neighbor (or KNN), the training data (which is well-labeled) is fed into the learner. When the test data is introduced to the learner, it compares both the data. k most correlated data is taken from training set. The majority of k is taken which serves as the new class for the test data [27].

III. Conclusion

This paper surveys various machine learning algorithms. Today each and every person is using machine learning knowingly or unknowingly. From getting a recommended product in online shopping to updating photos in social

networking sites. This paper gives an introduction to most of the popular machine learning algorithms.

REFERENCES

[1] W. Richert, L. P. Coelho, "Building Machine Learning Systems with Python", Packt Publishing Ltd., ISBN 978-1-78216-140-0
[2] M. Welling, "A First Encounter with Machine Learning"
[3] M. Bowles, "Machine Learning in Python: Essential Techniques for Predictive Analytics", John Wiley & Sons Inc., ISBN: 978-1-118-96174-2
[4] S.B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques", Informatica 31 (2007) 249-268
[5] L. Rokach, O. Maimon, "Top – Down Induction of Decision Trees Classifiers – A Survey", IEEE Transactions on Systems,
[6] D. Lowd, P. Domingos, "Naïve Bayes Models for Probability Estimation"
[7] https://webdocs.cs.ualberta.ca/~greiner/C-651/Homework2_Fall2008.html
[8] D. Meyer, "Support Vector Machines – The Interface to libsvm in package e1071", August 2015
[9] S. S. Shwartz, Y. Singer, N. Srebro, "Pegasos: Primal Estimated sub - Gradient Solver for SVM", Proceedings of the 24th International Conference on Machine Learning, Corvallis, OR, 2007
[10] http://www.simplilearn.com/what-is-machine-learning-and-why-itmatters-article
[11] P. Harrington, "Machine Learning in action", Manning PublicationsCo., Shelter Island, New York, 2012
[12] http://pypr.sourceforge.net/kmeans.html
[13] K. Alsabati, S. Ranaka, V. Singh, "An efficient k-means clustering algorithm", Electrical Engineering and Computer Science, 1997
[14] M. Andrecut, "Parallel GPU Implementation of Iterative PCA Algorithms", Institute of Biocomplexity and Informatics, University of Calgary, Canada, 2008
[15] X. Zhu, A. B. Goldberg, "Introduction to Semi – Supervised Learning", Synthesis Lectures on Artificial Intelligence and Machine Learning, 2009, Vol. 3, No. 1, Pages 1-130
[16] X. Zhu, "Semi-Supervised Learning Literature Survey", Computer Sciences, University of Wisconsin-Madison, No. 1530, 2005
[17] R. S. Sutton, "Introduction: The Challenge of Reinforcement Learning", Machine Learning, 8, Page 225-227, Kluwer Academic Publishers, Boston, 1992
[18] L. P. Kaelbing, M. L. Littman, A. W. Moore, "Reinforcement Learning: A Survey", Journal of Artificial Intelligence Research, 4, Page 237-285, 1996
[19] R. Caruana, "Multitask Learning", Machine Learning, 28, 41-75, Kluwer Academic Publishers, 1997
[20] D. Opitz, R. Maclin, "Popular Ensemble Methods: An Empirical Study", Journal of Artificial Intelligence Research, 11, Pages 169-198, 1999
[21] Z. H. Zhou, "Ensemble Learning", National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China
[22] https://en.wikipedia.org/wiki/Boosting_(machine_learning)
[23] https://en.wikipedia.org/wiki/Bootstrap_aggregating
[24] V. Sharma, S. Rai, A. Dev, "A Comprehensive Study of Artificial Neural Networks", International Journal of Advanced Research in Computer Science and Software Engineering, ISSN 2277128X, Volume 2, Issue 10, October 2012
[25] S. B. Hiregoudar, K. Manjunath, K. S. Patil, "A Survey: Research Summary on Neural Networks", International Journal of Research in Engineering and Technology, ISSN: 2319 1163, Volume 03, Special Issue 03, pages 385-389, May, 2014
[26] https://en.wikipedia.org/wiki/Instance-based_learning
[27] P. Harrington, "Machine Learning in Action", Manning Publications Co., Shelter Island, New York, ISBN 9781617290183, 2012
[28] J. M. Keller, M. R. Gray, J. A. Givens Jr., "A Fuzzy K-Nearest Neighbor Algorithm", IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-15, No. 4, August 1985
[29] Ayon Dey., "Machine Learning Algorithms: A Review", (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (3), 2016, 1174-1179.
[30] Kajaree Das, Rabi Narayan Behera, "A Survey on Machine Learning: Concept, Algorithms and Applications", (IJIRCCE) International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Issue 2, February 2017.

