A Review of Machine Learning and Deep Learning Applications

Pramila P. Shinde
Ph.D. Scholar at MPSTME, NMIMS&Assistant Professor at Shah and Anchor Kutchhi Engineering College, Mumbai.
pramila.shinde@sakec.ac.in

Dr. Seema Shah
Associate Professor & In-charge of B.Tech. Integrated Program at Mukesh Patel School of Technology Management and Engineering (MPSTME), NMIMS, Mumbai.
seema.shah@nmims.edu

Abstract—**Machine learning is one of the fields in the modern computing world.A plenty of research has been undertaken to make machines intelligent. Learning is a natural human behavior which has been made an essential aspect of the machines as well. There are various techniques devised for the same.Traditional machine learning algorithms have been applied in many application areas. Researchers have put many efforts to improve the accuracy of that machinelearning algorithms.Another dimension was given thought which leads to deep learning concept. Deep learning is a subset of machine learning. So far few applications of deep learning have been explored. This is definitely going to cater to solving issues in several new application domains, sub-domains using deep learning. A review of these past and future application domains, sub-domains, and applications of machine learning and deep learning are illustrated in this paper.**

Keywords—Machine Learning, Deep Learning, Frameworks.

I.   INTRODUCTION

Artificial Intelligence (AI) refers to making machines as intelligent as the human brain. In Computer Science, AI means the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its likelihood of with success achieving its goals. Informally, the term "artificial intelligence" is applied when a machine is able to perform functions that humans associate with other human minds, such as "learning" and "problem solving". Learning is a vital aspect of machines. Therefore, machine learningis a subfield of AI.Computer Scientists have taken efforts since the 1950s in the domain of machine learning. Since the last few decades tremendous efforts are made in the advancements of machine learning. This leads to higher expectations from machines. Deep learning is an attempt in this direction. It is a subset of machine learning. As the work in learning is put forward in many new areas and applicability of newer areas is always an undergoing task in the research community.

Field of Artificial Intelligence
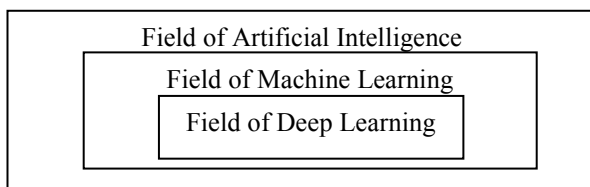Field of Machine Learning
Field of Deep Learning

Fig.1. depicts the earlier research from where deep learning has originated. Deep learning refers to deep artificial neural networks. Deep is the term which refers to a number of layers in a neural network. The deep network has more than one hidden layer whereas a shallow network has only one.

This paper presents a review of applications of machine learning and deep learning. Section II gives Overview of Machine Learning, Section III describes Machine Learning Applications, Section IV highlights Overview of Deep Learning, Section V briefs Deep Learning Applications and Section VI gives Inferences from the review of Applications of Machine Learning and Deep Learning which is followed by Conclusion.

II.   OVERVIEW OF MACHINE LEARNING

An overview of how machine learning is evolved is highlighted in this section. Intelligent Machinery was the term authored in the 1950s which acquainted the world with another area wherein machines were attempting to become intelligent as we human beings are. This was the initial move towards wandering into new era.In 1948, Turing and Champernowne found 'paper and pencil' chess. It was the world's first chess playing computer program. The program was formulated with pencil and paper, the estimations being performed physically by Turing and Champernowne themselves – each move would take them thirty minutes or more to ascertain. Dietrich Prinz composed program mate-in-two moves chess machine in 1951. The program presented a piece-list in conjunction with an installed post box board portrayal, yet 10*10, since a knight move was made out of two single step moves. By having a ply-indexed array of piece-list-index, direction- and step-counter move generation was done. Christopher Strachey programs first Draughts (Checkers) algorithm in 1952. The program could play an entire session of Draughts at a reasonable speed [1].

The first artificial intelligence program to incorporate learning, written by Anthony Oettinger was called as "response learning programme" and "shopping programme" in the year 1951. The shopping program reproduced the conduct of a little kid sent on a shopping visit. This was the main endeavor towards learning machines. In the year1955, Arthur Samuel adds learning to his Draughtsalgorithm. It is the

principal machine learning framework that got open acknowledgment. Itisdraughts-playing program that human opponents described as "tricky but beatable".

In 1943, The McCulloch-Pitts Model of Neuron worked by inputting either a 1 or 0 for each of the inputs, where 1 represented true and 0 false. Likewise, the threshold was given a real value, say 1, which would allow for a 0 or 1 output if the threshold was met or exceeded. In 1949, Donald Hebb proposed that when two neurons fire together the connection between the neurons is strengthened. Moreover, this action is one of the fundamental operations necessary for learning and memory. Frank Rosenblatt, utilizing the McCulloch-Pitts neuron and the discoveries of Hebb, went ahead to build up the primary perceptron in 1957 which is known as perceptron learning rule In 1962 Rosenblatt proves the perceptron convergence theorem. After three years, Bernard Widrowen graved Delta Learning rule. It is used for perceptron training which is also called as the Least Square problem. A good linear classifier is created by combining these two. However, Marvin Minsky in 1969 proposed the XOR problem. He also showed the inability of Perceptrons in such linearly inseparable data distributions. Thereafter, NN researches were dormant up till 1980s[2].

Seppo Linnainmaa in 1970 proposed "reverse mode of automatic differentiation". It is Backpropagation(BP) algorithm. Later, Paul Werbos suggested Multi-Layer Perceptron (MLP) with NN specific Backpropagation(BP) algorithm in 1981. In 1985–1986 NN researchers (Rumelhart, Hinton, Williams—Hetch, Nielsen) successively presented the concept of MLP with practical BP training.

In 1986 David Rumelhart and James McClelland present the multilayer perceptron. Hopfield Network (Or "Hopfield model") is a kind of neural network proposed by John Hopfield in the early 1980s. The Hopfield network has no special input or output neurons, but all are both input and output, and are connected to all others in both directions (with equal weights in the two directions). Hinton and Sejnowski (1986) devised Boltzmann machines by combining Hopfield networks and simulated annealing. It is a fully connected two-layer neural network.

The major machine learning achievement was Support Vector Machines (Networks) (SVM), proposed by Vapnik and Cortes in 1995 with solid hypothetical standing and exactoutcomes. That was the time isolating the machine learning group into two groups as NN or SVM researchers. Freund and Schapire in 1997 developed boosted ensemble of weak classifiers called Adaboost. AdaBoost was explored by Breiman in 2001 that ensemble multiple decision trees where each of them is created by a random subset of instances and each node is selected from a random subset of features. It is known as Random Forests (RF). AdaBoost algorithm shows

weakness to over-fitting and outlier instances in the data whereas RF is more robust model [1][3].

As we come nearer today, a new era of NN called Deep Learning has evolved. The third ascent of NN has started generally in 2005 with the conjunction of a wide range of discoveries from over a significant time span by late researchers Hinton, Andrew Ng, LeCun, Bengio, and other researchers.

The machine learning algorithms which are widely in use are Linear Classifier, Logistic Regression, Naïve Bayes (NB), Bayesian Network, Support Vector Machines (SVM), Decision Tree, Random Forest, AdaBoost, Bootstrapped Aggregation (Bagging), k-Nearest Neighbour (k-NN) and Artificial Neural Network (ANN).

There is a wide range of open source machine learning frameworks available in the market, which enables machine learning engineers to create, implement and maintain machine learning systems, generate new projects and create new impactful machine learning systems.Several of machine learning frameworks and its version available as of 2017 are as follows: Apache Singa (1.0), Shogun (4.1.0), Apache Mahout (0.12.2), Apache Spark MLib (2.0.1), TensorFlow (0.10.0), Oryx 2 (2.2.1), Accord.NET (3.1.0) and Amazon Machine Learning (2017). Using above frameworks, any machine learning application can be implemented.

III. MACHINE LEARNING APPLICATIONS

While surveying through we arrived at various application domains and sub-domains of machine learning applications.

- The different application domains are Computer vision, prediction, semantic analysis, natural language processing and information retrieval.
- Computer Vision: Object recognition, object detection, and object processing are sub-domains in computer vision domain.
- Prediction: The various sub-domains here are classification, analysis, and recommendation. Text classification, document classification, image analysis, medical diagnosis, prediction of network intrusion detection and predicting denial of service attack have been successfully implemented using machine learning.
- Semantic Analysis and Natural Language Processing and Information Retrieval: Semantic analysis is the process of relating syntactic structures from paragraphs, sentences, words to the level of writing as a whole. Natural language processing is how to program computers to correctly process natural language data. Information retrieval is the science of searching for information in a document, searching for documents and searching for metadata that describe the data and for databases of sounds and images. These are three domains in which machine learning techniques have been explored in the past.

An attempt to classify all machine learning applications found in the past is summarized here. The following Fig. 2 depicts the same as application domains at first level, sub-domains at second level and few applications at last level which have been implemented by researchers.
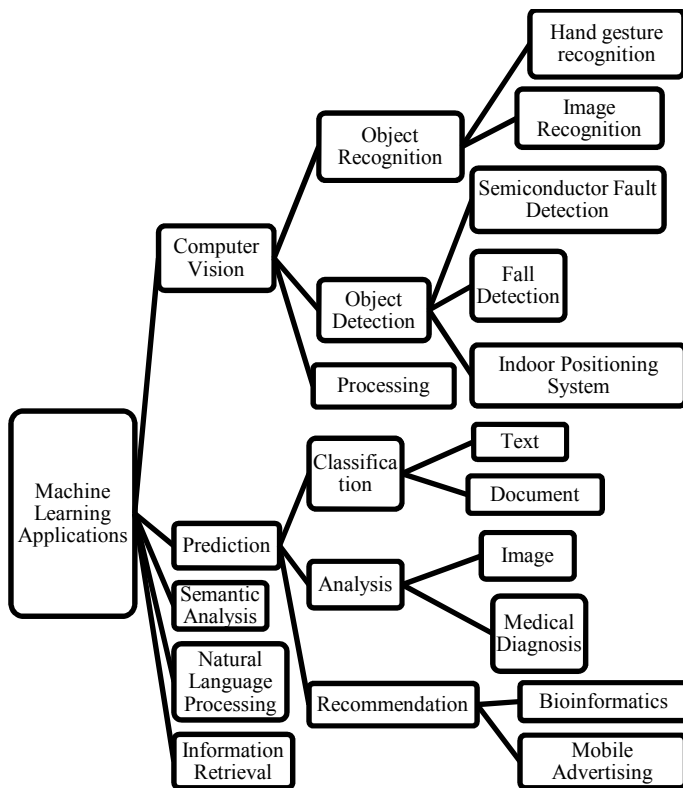

Fig. 2. Machine Learning Applications

Let us briefly look at the machine learning applications.Hand gesture recognition for human-computer interaction is implemented using k-Nearest Neighbour (k-NN), Naïve Bayes (NB), Artificial Neural Network (ANN) and Support Vector Machines (SVM) which are analyzed with parameters learning rate, momentum rate, training cycles, hidden layers in the network [4].

The classification application such as network intrusion detection system using BayesNet, Logistic, IBk, JRip, PART,J48,Random Forest, RandomTree, REPTreecompares parameters Receiver Operating Characteristics (ROC) curve, sensitivity, specificity, precision, accuracy, kappa, minimum absolute error,F1 score,False Positive Rate(FPR),Negative Predictive rate(NPV),False Discovery rate(FDR),training time(Seconds) [5].

Object detection kind of machine learning application is indoor positioning systemswhich use decision tree, naïve bayes, Bayesian Network (BN), k-Nearest Neighbour,Sequential Minimal Optimization(SMO),AdaBoost, Bagging with comparison parameters as computational time, and accuracy [6].

Another classification example is classifying DDoS attacks using Naïve Bayes, RBF network, Multi Layer Perceptron, Bayesnet, IBK, J48, Voting,Bagging+Random Forest, Random Forest, Adaboost+Random Forest with output parameters as accuracy, False Negative (FN) rate, False Positive (FP) rate, precision, recall [7].

Fault detection modelsare done usingk-NN,SVMtechniques with evaluating parameters such as Geometric Mean (G-mean), Harmonic Mean (F-measure)are fully described [8].

Best Classifier algorithm from Naive Bayes, IBK, J48, Adaboost, LogitBoost, PART, Random Forest, Bagging, and SMOfor a given dataset is found with the help of parameters Number of attributes, Number of instances, Number of classes, Kurtosis, Skewness, Maximum Probability, and Entropy [9].

Intrusion Detection is implemented usingSVM, J48, Naive Bayes, Decision Table techniques with parameters True Positive Rate(TPR),False Positive Rate (FPR), Precisionto see which technique gives better results[10].

Breast Cancer Detection and Diagnosisis object detection topic using SVM,BN,Random Forest techniques with output parameters such as precision, recall, and Area under ROC valuesis implemented [11].

Text Classificationis done using Naïve Bayes, Linear Classifier, SVM, ANN techniqueswith only one output parameter as accuracy for determining which technique is best [12].

Example of Prediction is the Education System usingLinear Regression, Logistic Regression, Decision Tree, Naive Bayes, and Support Vector Machine with output parameters Learning skills and Psychological factors [13].The results from running mentioned algorithms have been analyzed with respect to output parameters as mentioned in above paragraphs. Accuracy is one of the common parameters.

IV. OVERVIEW OF DEEP LEARNING

Deep learning is a subset of machine learning. It is a neural network with a large number of layers and parameters. Most deep learning methods use neural network architectures. Therefore it is also referred to as deep neural networks.

In short, deep learning uses a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. The lower layers close to the data input learn simple features, while higher layers learn more complex features derived from lower layer features. The architecture forms a hierarchical and powerful feature representation. It means that deep learning is suited for analyzing and extracting useful knowledge from both large huge amounts of data and data collected from different sources [14].

2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)

The NN researchers have taken efforts to continuously add developments to the field. To begin with, Self-organizing neural networks (1980) are used to cluster input patterns into groups of similar patterns. They're called "maps" since they assume a topological structure among their cluster units; effectively mapping weights to input data. The Kohonen network introduces the concepts of self-organization and unsupervised learning.

Kunihiko Fukushima proposed the neocognitron in the 1980s which is a hierarchical, multilayered artificial neural network. It provides solution to handwritten character recognition and other pattern recognition tasks. Later on, Convolutional neural networks were developed.

A Restricted Boltzmann Machine (RBM) is a generative stochastic artificial neural network that can learn a probability distribution over its set of inputs. It became popular after Geoffrey Hinton and collaborators developed fast learning algorithms for them in the mid-2000.

The Recurrent Neural Networks (RNN) is useful for recovering a stored pattern from a corrupted version and is the predecessors of Boltzmann machines and auto-encoders. Michael Jordan invented Jordan networks in 1986 which is an early architecture for supervised learning on sequences.

The first Convolutional Neural Network (CNN) — LeNet architecture was first introduced by LeCunet al. in their 1998 paper, Gradient-Based Learning applied to document recognition. It was used mainly for OCR and character recognition in documents.

The Bidirectional Recurrent Neural Networks (BRNN) was introduced by Mike Schuster and Kuldip Paliwal in 1997. It works by training network simultaneously in positive and negative time direction.

Long Short-Term Memory (LSTM) was presented by Hochreiter and Schmidhuber in 1997. LSTM can learn to bridge minimal time lags in excess of 1000 discrete time steps by enforcing constant error flow through "constant error carrousels" within special units. Multiplicative gate units learn to open and close access to the constant error flow. LSTM is local in space and time; its computational complexity per time step and weight is O (1).

Deep Belief Networks was followed by Deep Boltzmann machines in the same year 2006. Dropout neural networks were developed in the year 2012 by Hinton. Generative Adversarial Networks (GANs) are deep neural net architectures which were invented in the year 2014 by Goodfellow. Capsule neural networks are recent advancements in deep learning by Hinton.

The main reasons for the popularity of deep learning today are, drastically increased chip processing abilities (e.g., GPU units), the significant low cost of computing hardware, and recent advances in machine learning and signal/information processing research [15].

The typical deep learning techniques are Autoencoder (AE), Deep Belief Network (DBN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Recursive Neural Network and Direct Deep Reinforcement Learning.

There are various deep learning frameworks which researchers can use for implementing any deep learning technique. Various deep learning frameworks and its version available as of 2017 are as follows:Theano by MILA (2017), TensorFlow by Google (0.10.0), PyTorch by Facebook (2017), Microsoft Cognitive Toolkit by Microsoft (2.0), Caffe2 by Facebook (2017), MXNet by Apache (1.0.0), Deeplearning4jOpen-Source (0.9.1), H2o.aiOpen-Source (Tutte 3.10.2.2), Spark by Apache (2.2.1). The framework can be selected based on deep learning application, technique, and platform for implementation.

V. DEEP LEARNING APPLICATIONS

Few of newer and recent application developments of deep learning are elaborated in examples below:

A. One example of an application of deep learning in Big Data is Microsoft speech recognition (MAVIS). Using deep learning enables searching of audios and video files through human voices and speeches [16].

B. Deep learning on Big Data environment is used by Google for image search service. They used deep learning for understanding images so that it can be used for image annotation and tagging that is further useful in image search engines and image retrieval as well as image indexing [16].

C. In 2016, Google's AlphaGo program defeated Lee Sedol in Go competition, which showed that deep learning had a strong learning ability.

D. Google's Deep Dream is software which can not only classify images but also generate strange and artificial paintings based on its own knowledge.

E. Facebook announced a new artificial intelligence system named Deep Text. It is a deep learning-based text understanding engine which can classify massive amounts of data, provide corresponding services for identifying users chatting messages and clean up spam messages.

- Computer vision, prediction, semantic analysis, natural language processing, information retrieval and customer relationship management are the application domains for deep learning.

- Computer vision has object recognition, object detection and processing as sub-domains.It includes automatic speech recognition, image recognition, speech and audio processing and visual art processing

2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)

which newer applications are being explored using deep learning.

- Prediction: The different sub-domains here are classification, analysis, and recommendation.Drug discovery and toxicology, Bioinformatics, mobile advertising are newer applications being developed using deep learning.

- Semantic Analysis and Natural Language Processing and Information Retrieval: These are three areas in which machine learning techniques have been explored in the past but additionally now researchers are applying deep learning techniques.

- Customer Relationship Management: Given customer's history, data analysis is done which is used to enhance business relationships. In this case, deep learning can be useful. These above mentioned deep learning applications have been presented in Fig. 3 below.
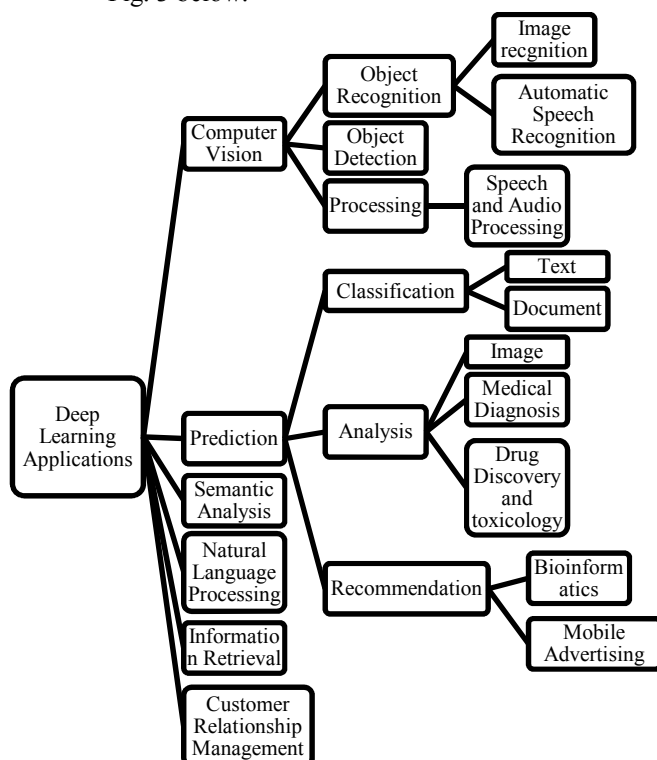
Fig.3. Deep learning applications

Let us briefly review these applications. Computer vision (such as object detection, object tracking, and image segmentation) and natural language processing application is presented using deep learning techniques such as Autoencoder, Deep Belief Network, Convolutional Neural Network and Recurrent Neural Network in this paper [17].

Medical domain topics such as translational bioinformatics,

Medical imaging, pervasive sensing, medical informatics, and public health can be implemented using the Deep Belief Network and the Deep Boltzmann Machine (DBM) techniques [18].

A prediction domain topic such as Financial Signal Representation and Trading can be very well be implemented usingDirect Deep Reinforcement Learning technique in this paper [19].

Microscopy Image Analysis is often done by exploitation Deep Neural Network as can be seen from this paper [20].

Different application areas such as object recognition, speech and audio processing and tagging, information retrieval, natural language processingcan be addressed usingDeep Belief Network (DBN), Deep Boltzmann Machines (DBM) andDeep Stacking Networks (DSN) algorithms [21].

Applications of deep learning are to be explored in newer areas such as medical imaging, deep learning creating sound, deep learning doing art, computer hallucinations, robots, predictions, computer games, self-driving cars, and Big Data. Existing areas also have few problems which can be overcome by using deep learning such as automatic colorization, automatic machine translation, automatic text generation, automatic handwriting generation, image recognition, automatic image caption generation, advertising (create data-driven predictive advertising, real-time bidding-RTB for their ads, precisely targeted display advertising and more), predicting earthquakes and so on.

VI. INFERENCES FROM REVIEW OF APPLICATIONS OF MACHINE LEARNING AND DEEP LEARNING

Let us infer from above study the common application areas and differences. The common areas of applications of machine learning and deep learning are Computer vision, prediction, semantic analysis and natural language processing. Customer relationship management is one of the newer areas as application of deep learning.

The primary reason why deep learning is suited for newer areas of applications is data dependencies, GPU hardware, and feature engineering. Data dependencies term is to refer to deep learning algorithms work well on a huge amount of data. GPU is a high-end machine which stands for Graphics Processing Unit. The distinctive part of deep learning when compared to machine learning is the ability to learn high-level features from data called as feature engineering. Therefore, there could be many more areas of applications of deep learning which will be seen in forthcoming years.

VII. CONCLUSION

This paper introduces the need to study machine learning and deep learning. It has presented machine learning and deep learning evolution along with their applications which have been explored by the researchers in the last few decades. For developing any application in machine learning or deep learning a range of frameworks are availablewhich are discussed. There are many new areas of applications for deep learning. There is a still lot of scope for digging deep into what could be application area of deep learning.

With this review of applications, now we can explore any one of the newer areas of application of deep learning which will yield better results and will add on to the ongoing research in this field. There is even scope for evolving new architectures for deep learning as research is still going on in the early stage. Apart from this enhancements can be done in analysis and prediction sub-domain.

REFERENCES

[1] S.Angra and S.Ahuja "Machine Learning and its application" in International Conference on Big Data Analytics and Computational Intelligence (ICBDAC),2017.

[2] Mathews Chibuluma and JosephatKalezhi "Application of a modified perceptron learning algorithm to monitoring and control"in2017 IEEE PES PowerAfrica.

[3] Amanpreet Singh and Narina Thakur" A review of supervised machine learning algorithms" in 3rd International Conference on Computing for Sustainable Global Development (INDIACom) 2016.

[4] P. Trigueiros, F. Ribeiro, and L. P. Reis, "A comparison of machine learning algorithms applied to hand gesture recognition."

[5] S. Choudhury and A. Bhowal, "Comparative analysis of machine learning algorithms along with classifiers for network intrusion detection," in 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 2015.

[6] S. Bozkurt, G. Elibol, S. Gunal, and U. Yayan, "A Comparative Study on Machine Learning Algorithms for Indoor Positioning."

[7] R. R. R. Robinson and C. Thomas, "Ranking of machine learning algorithms based on the performance in classifying DDoS attacks," in IEEE Recent Advances in Intelligent Computational Systems (RAICS), 2015.

[8] T. Lee, K. B. Lee, and C. O. Kim, "Performance of Machine Learning Algorithms for Class-Imbalanced Process Fault Detection Problems," IEEE Trans. Semicond. Manuf., 2016.

[9] N. Pise and P. Kulkarni, "Algorithm Selection for Classification Problems," SAI Comput. Conf., 2016.

[10] T. Mehmood and H. B. MdRais, "Machine Learning Algorithms In Context Of Intrusion Detection," vol. 369, 2016.

[11] D. Bazazeh and R. Shubair, "Comparative Study of Machine Learning Algorithms for Breast Cancer Detection and Diagnosis."

[12] S. ZamanMishu, "Performance Analysis of Supervised Machine Learning Algorithms for Text Classification."

[13] R. R. Halde, "Application of Machine Learning algorithms for betterment in Education system."

[14] Lei Zhang, Shuai Wang, Bing Liu, "Deep Learning for Sentiment Analysis: A Survey", National Science Foundation (NSF), and by Huawei Technologies Co. Ltd., 2017.

[15] Markoff, J., "Scientists SeePromisein Deep-LearningPrograms",NewYork Times, November 23, 2012.

[16] M. Gheisari, G. Wang, and M. Z. A. Bhuiyan, "A Survey on Deep Learning in Big Data," in 22017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), 2017.

[17] X. Du, Y. Cai, S. Wang, and L. Zhang, "Overview of deep learning," in Proceedings - 2016 31st Youth Academic Annual Conference of Chinese Association of Automation, YAC 2016, 2017.

[18] D. Ravi et al., "Deep Learning for Health Informatics," IEEE J. Biomed. Heal. Informatics, 2017.

[19] Y. Deng, F. Bao, Y. Kong, Z. Ren, and Q. Dai, "Deep Direct Reinforcement Learning for Financial Signal Representation and Trading," IEEE Trans. Neural Networks Learn. Syst., 2017.

[20] F. Xing, Y. Xie, H. Su, F. Liu, and L. Yang, "Deep Learning in Microscopy Image Analysis: A Survey", IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, 2017.

[21] N. M. Elaraby, M. Elmogy, and S. Barakat, "Deep Learning: Effective Tool for Big Data Analytics", International Journal of Computer Science Engineering (IJCSE), Sep 2016.