

Deep Machine Learning and Neural Networks: An Overview

Chandrabhas Mishra¹ and Dharmendra Lal Gupta²

¹*M.Tech in Computer Science & Engineering KNIT Sultanpur, India -228118*

²*Associate Professor in Computer Science & Engineering Dept. KNIT Sultanpur, India – 228118*

¹*cmrj88@gmail.com, ²dlgupta2002@gmail.com*

Abstract

Deep learning is a technique of machine learning in artificial intelligence area. Deep learning is a refined "machine learning" algorithm that surpasses a considerable lot of its forerunners in its capacity to perceive syllables and pictures. As of now Deep learning is a greatly dynamic examination territory in machine learning and example acknowledgment society. It has increased colossal triumphs in an expansive zone of utilizations, for example, speech recognition, computer vision and natural language processing and numerous industry items. Neural networks are used to implement the machine learning or to design intelligent machines. In this paper thorough survey to all machine learning paradigms and application areas of deep machine learning and different types of neural networks with applications are discussed.

Keywords: *Feature Representation, Machine learning (ML), Neural Nets Models, Deep Neural Network (DNN) Convolutional Neural Networks (CNNs) and Deep Belief Networks (DBNs), Automatic Speech Recognition (ASR) and Artificial Neural Network (ANN)*

1. Introduction

Machine learning:

Learning is a process in which association of events with consequences is done. Thus basically learning is a way to substantiate the cause and effect principle. The science of designing the intelligent machine is referred to as machine learning and the tool used to design such intelligent machine is neural networks. Neural network may considered as black-box which gives some desired output for the given input. It is achieved through process called training.

In contrast to most conventional learning methods, which are considered utilizing shallow-organized learning designs, deep learning suggests machine learning methods that utilized supervised and/or unsupervised procedures to consequently learn various leveled representations in deep models for classification. Inspired by natural perceptions on human cerebrum systems for processing of natural signals, deep learning has attracted much attention from the academic community in recent years due to its state-of-the-art performance in many research domains such as speech recognition, collaborative filtering, and computer vision. Deep learning has additionally been effectively connected in industry items that exploit the expansive volume of advanced information. Companies like Google, Apple, and Facebook, who collect and analyses massive amounts of data on a daily basis, have been aggressively pushing forward deep learning related projects. For example, Apple's Siri, the virtual personal assistant in iPhones, offers a wide assortment of administrations including climate reports, sport news, answers to user's questions, and reminders *etc.* by utilizing deep learning and more and more data collected by Apple services. Google applies deep learning algorithms to monstrous pieces of chaotic information got from the Internet for Google's translator.

Deep learning refers to a class of ML techniques, where numerous layers of data processing stages in various leveled architectures are exploited for unsupervised feature learning and for pattern classification. It is in the convergences among the research area of neural system, optimization, graphical modelling, signal processing, and pattern recognition. Two vital explanations behind the prominence of deep learning today are the fundamentally brought down expense of processing equipment and the drastically increased chip processing abilities *e.g.*, GPU units. Success of deep learning in various applications including computer vision, phonetic recognition, voice search, spontaneous speech recognition, speech and image feature coding, semantic utterance classification, hand-writing recognition, audio processing, information retrieval, and robotics and many artificial intelligence area. Before going to deal with different machine learning paradigm in detail, a brief classification is shown in Figure 1. We use four key attribute to classify the machine learning paradigm

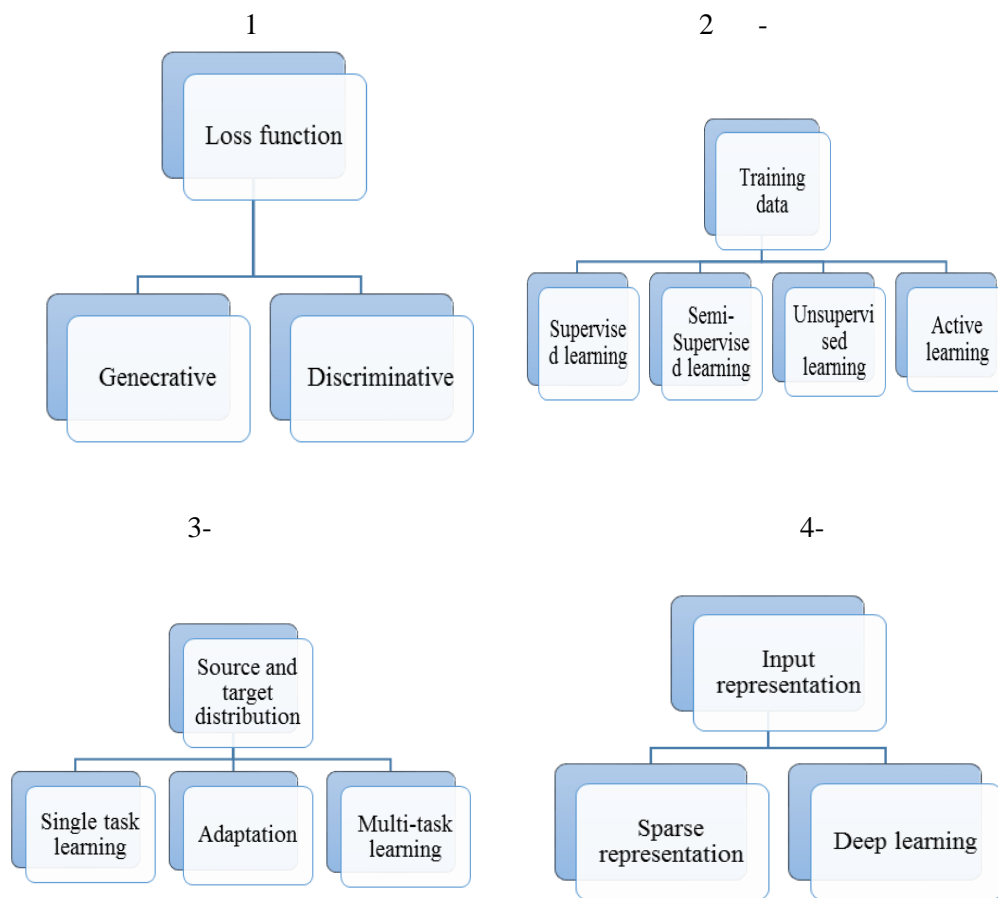


Figure 1. Classification of Machine Learning Paradigm

1.1. Generative Learning

Generative learning and discriminative learning are the two most prevalent, antagonistically paired ML paradigms developed and deployed in ASR (Automatic speech recognition). There are two key factors that distinguish generative learning from discriminative learning: the nature of the model (and hence the decision function) and the loss function (*i.e.*, the core term in the training objective). Briefly speaking, generative learning consists of

- Using a generative model, and
- Adopting a training objective function based on the joint likelihood loss defined on the generative model.

Discriminative learning, on the other hand, requires either

- Using a discriminative model, or
- Applying a discriminative training objective function to a generative model.

Here we will discuss generative vs. discriminative learning from both the model and loss function perspectives. While historically there has been a strong association between a model and the loss function chosen to train the model, there has been no necessary pairing of these two component in the literature.

1.2. Discriminative Learning

As discussed earlier, the paradigm of discriminative learning involves either using a discriminative model or applying discriminative training to a generative model. We first give a general discourse of the discriminative models and of the discriminative loss functions utilized as a part of training, followed by an overview of the use of discriminative learning in ASR applications.

Table 1. List of Symbols and Their Meanings Used In Equations

<u>Symbol</u>	<u>Meaning</u>
\mathcal{X}	Space of input vectors
\mathcal{Y}	Set of output labels
$p(\mathbf{x}, y)$	Joint distribution $p(\mathbf{X} = \mathbf{x}, Y = y)$
\mathcal{F}	Space of decision functions $f : \mathcal{X} \rightarrow \mathcal{Y}$
$f(\mathbf{x}; \lambda)$	Decision function
$d_y(\mathbf{x}; \lambda)$	Discriminant function
λ	Model or decision function parameters
$L(f(\mathbf{x}), y)$	Loss function
$E_{p(\mathbf{x}, y)}[\cdot]$	Expectation $E_{(\mathbf{x}, y) \sim p(\mathbf{x}, y)}[\cdot]$
\mathcal{D}	Training data $(\mathbf{x}^{(i)}, \mathbf{y}^{(i)})_{i=1}^m$

Models:

Discriminative models make direct use of the conditional relation of labels given input vectors. One of such major models are referred to as *Bayesian Minimum Risk* (BMR) classifiers. Shown in equation 1.

$$f(\mathbf{x}; \lambda) = - \arg \min_{y'} \sum_y \Delta(y', y) p(y|\mathbf{x}; \lambda) \quad 1$$

Loss Functions:

The first group of loss functions are based on probabilistic models, while the second group on the notion of *margin*.

1) *Probability-Based Loss*: Similar to the joint likelihood loss, *conditional likelihood loss* is a probability-based loss function but is defined upon the conditional relation of class labels given input features: . Shown in equation 2.

$$L(f(\mathbf{x}), y) = - \ln p(y|\mathbf{x}; \lambda)$$

This loss function is strongly tied to probabilistic discriminative models such as conditional log linear models and MLPs, while they can be applied to generative models as well. Moreover, conditional likelihood loss can be naturally extended to predicting structure output. For example, when applying to Markov random fields, we obtain the training objective of conditional random fields (CRFs): by equation 3.

$$p(\mathbf{y}|\mathbf{x}; \lambda) = \frac{\exp \lambda \cdot f(\mathbf{y}, \mathbf{x})}{Z_{\lambda}(\mathbf{x})} \quad \text{In various ML}$$

techniques, one often calls the training method using the conditional likelihood loss above as simply maximal likelihood estimation (MLE). A generalization of conditional likelihood loss is Minimum Bayes Risk training. This is consistent with the criterion of MBR classifiers described in the previous subsection. The loss function of (MBR) in training is given by equation 4.

$$L(f(\mathbf{x}), \mathbf{y}) = -\ln \sum_{\mathbf{y}'} \Delta(\mathbf{y}', \mathbf{y}) p(\mathbf{y}'|\mathbf{x}; \lambda) \quad 4$$

1.3. Supervised and Unsupervised Learning

Supervised and unsupervised learning are the techniques of machine learning.

1.3.1. Supervised Learning

In supervised learning, the training set consists of pairs of inputs and outputs drawn from a joint distribution.

$$\mathcal{D} = \{(\mathbf{x}^{(i)}, y^{(i)}) | (\mathbf{x}^{(i)}, y^{(i)}) \sim p(\mathbf{x}, y)\}_{i=1}^n \quad \text{Using}$$

notations introduced by equation 5.

The learning objective is again empirical risk minimization with regularization, *i.e.* where both input data and the corresponding output labels are provided. Notice that there may exist multiple levels of label variables. In this case, we should differentiate between the *fully supervised* case, where labels of all levels are known, the *partially supervised* case, where labels at certain levels are missing.

1.3.2. Unsupervised Learning

In ML, unsupervised learning is generally refers to learning with the input data only. This learning paradigm often aims at building representations of the input that can be used for prediction, decision making or classification, and data compression. For example, density estimation, clustering, principle component analysis and independent component analysis are all important forms of unsupervised learning. Use of vector quantization (VQ) to provide discrete inputs to ASR is one early successful application of unsupervised learning to ASR [1]. More recently, unsupervised learning has been developed as a component of staged hybrid generative-discriminative paradigm in ML. Learning sparse speech representations, can likewise be viewed as unsupervised feature learning or learning feature representations in absence of classification labels.

1.4. Passive and Active Learning

The preceding overview of generative and discriminative ML paradigms uses the attributes of loss and decision functions to organize a multitude of ML techniques. In this section, we use a different set of attributes, namely the nature of the training data in

relation to their class labels. Depending on the way that training samples are labelled or otherwise, we can classify many existing ML techniques into several separate paradigms. Supervised learning assumes that all training samples are labelled, while unsupervised learning assumes none. Semi-supervised learning, as the name suggests, assumes that both labelled and unlabeled training samples are available. Supervised, unsupervised and semi-supervised learning are typically referred to under the *passive learning* setting, where labelled training samples are generated randomly according to an unknown probability distribution. In contrast, *active learning* is a setting where the learner can intelligently choose which samples to label, which we will discuss at the end of this section. In this section, we concentrate mainly on semi-supervised and active learning paradigms. This is because supervised learning is reasonably well understood and unsupervised learning does not directly aim at predicting outputs from inputs.

1.4.1. Semi-Supervised Learning

The semi-supervised learning paradigm is of special significance in both theory and applications. In many ML applications, unlabeled data is discarded but labelling is expensive and time-consuming. It is possible and often helpful to leverage information from unlabeled data to influence learning. Semi-supervised learning is targeted at precisely this type of scenario, and it assumes the availability of both labelled and unlabeled data, *i.e.* given by equation 6.

$$\begin{aligned} \bullet \mathcal{D} &= \{(\mathbf{x}^{(i)}, y^{(i)}) | (\mathbf{x}^{(i)}, y^{(i)}) \sim p(\mathbf{x}, y)\}_{i=1}^m & 6 \\ \bullet \mathcal{U} &= \{\mathbf{x}^{(i)} | \mathbf{x}^{(i)} \sim p(\mathbf{x})\}_{i=m+1}^{m+n} \end{aligned}$$

Here we categorized semi-supervised learning methods based on their *inductive* or *transductive* nature. The key difference between inductive and transductive learning is the outcome of learning process. In the former setting, the goal is to find a decision function that not only correctly classifies training set samples, but also generalizes to any future sample. In contrast, transductive learning aims at directly predicting the output labels of a test set, without the need of generalizing to other samples. In this regard, the direct outcome of transductive semi-supervised learning is a set of labels instead of a decision function. All learning paradigms we have presented are inductive in nature.

1.4.2. Active Learning

Active learning is a similar setting to semi-supervised learning in that, in addition to a small amount of labelled data, there is a large amount of unlabeled data available. The goal of active learning, however, is to query the most informative set of inputs to be labelled, hoping to improve classification performance with the minimum number of queries. That is, in active learning, the learner may play an *active* role in deciding the data set rather than it be passively given. The key thought behind active learning is that a ML calculation can accomplish more noteworthy execution, *e.g.*, higher classification accuracy, with fewer training labels if it is allowed to choose the subset of data that have labels. An active learner might posture questions ordinarily as unlabeled information cases to be named regularly by a human. For this reason, it is sometimes called *query learning*. Active learning is all around inspired in numerous present day ML issues where unlabeled information might be copious or effortlessly gotten yet names are troublesome tedious, or expensive to obtain. Broadly, active learning comes in two forms: batch active learning, where a subset of data is chosen, *a priori* in a batch to be labelled. The labels of the instances in the batch chosen to be labelled may not, under this approach, influence other instances to be selected since all instances are chosen at once. In *online active learning*, on the other hand, instances are chosen one-by-one, and the true labels of all

previously labelled instances may be used to select other instances to be labelled. For this reason, online active learning is sometimes considered more powerful.

Artificial neural network:

An artificial neural network is an interconnected group of nodes, distantly related to the wide network of neurons in a brain shown in Figure 2. Here, every circular node shows an artificial neuron and an arrow shows to an association from the output of one neuron to the input of another should (ideally) be able to handle this. Artificial neural network consist of three type of layer namely input layer, hidden layer and output layer. Hidden layer is connected between input and output layer.

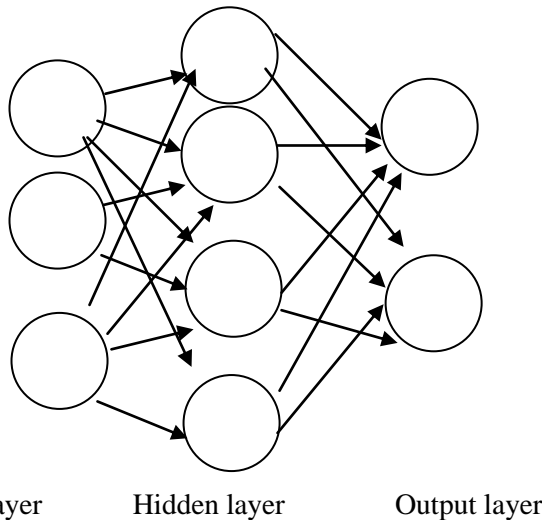


Figure 2. Artificial Neural Network Architecture

Convolutional Neural Networks:

CNNs are multi-layer neural networks shown in Figure 3 [2]. Convolutional neural network is basically designed for two-dimensional data, such as images and videos. Prior work in time-delay neural systems (TDNN) affected the CNNs, which diminish learning calculation necessities by sharing weights in a fleeting measurement and are planned for speech and time-series processing. The primary genuinely successful deep learning methodology is CNN, where numerous layers of an order were effectively trained in a robust manner. A CNN is a decision of engineering that influences spatial and temporal connections to lessen the quantity of parameters which must be scholarly and consequently improves upon general feed-forward back propagation training. CNNs were proposed as a deep learning system that is spurred by negligible information pre-processing necessities. In CNNs, little partitions of the picture are dealt with as inputs to the lowest layer of the various leveled structure or hierarchical structure. Information generally spreads through the diverse layers of the system whereby at every layer digital filtering is applied, keeping in mind the end goal to get remarkable features of the information observed.

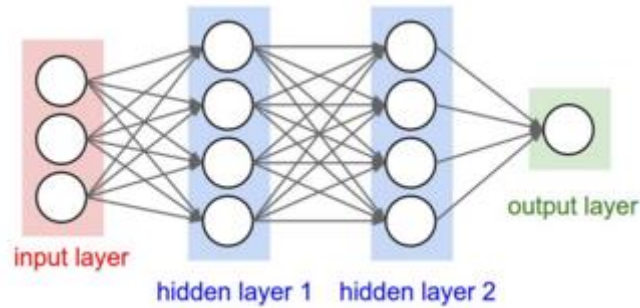


Figure 3. Convolutional Neural Network (CNN) Architecture

Deep Belief Networks:

DBNs are made out of a few layers of Restricted Boltzmann Machines, a type of neural network shown in Figure 4 [3]. These networks are “restricted” to a single visible layer and single hidden layer, where associations are formed between the layers (units inside a layer are not associated). The higher-order data correlations are captured by training the hidden units that are observed at the visible units. Initially, aside from the top two layers, which form an associative memory, directed top-down generative weights are used to connect the layers of a DBN. RBMs are appealing as a building piece, over more traditional and deeply layered sigmoid belief networks, due to their ease of learning these connection weights. The underlying pre-training happens in an unsupervised avaricious layer-by-layer way to get generative weights, empowered by what Hinton has termed contrastive disparity. Amid this training stage, a vector v is introduced to the visible units that forward values to the hidden units. Going in reverse, the visible unit inputs are then stochastically found with an end goal to reproduce the original input. At long last, these new visible neuron activations are sent such that one stage recreation hidden unit initiations can be achieved. Performing these forward and backward steps is a process known as Gibbs sampling, and the difference in the correlation of the hidden activations and visible inputs forms the basis for a weight update. Training time is significantly reduced as it can be shown that only a single step is needed to approximate maximum likelihood learning. Each layer added to the network can improve the log probability of the training data, which we can think of as increasing true representational power of network. This meaningful expansion, in conjunction with the utilization of unlabeled data, is a critical component in any deep learning application.

At the top two layers, the weights are tied together, such that the output of the lower layers provides a reference clue or link for the top layer to “associate” with its memory contents. We often encounter problems where discriminative performance is of ultimate concern, *e.g.* in classification tasks. A DBN may be fine-tuned after pre-training for improved discriminative performance by utilizing labelled data through back-propagation. At this point, a set of labels is attached to the top layer (expanding the associative memory) to clarify category boundaries in the network through which a new set of bottom-up, recognition weights are learned. It has been shown that such networks often perform better than those trained exclusively with backpropagation. This may be intuitively explained by the fact that back-propagation

For DBNs is only required to perform a local search on the weight (parameter) space, speeding training and convergence time in relation to traditional feed-forward neural networks

DBNs are stacks of restricted Boltzmann machines forming deep (multi-layer) architecture.

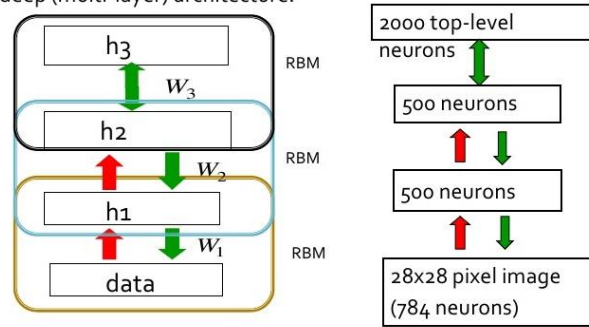


Figure 4. Deep Belief Neural Network (DBNN) Architecture

2. Research Methods

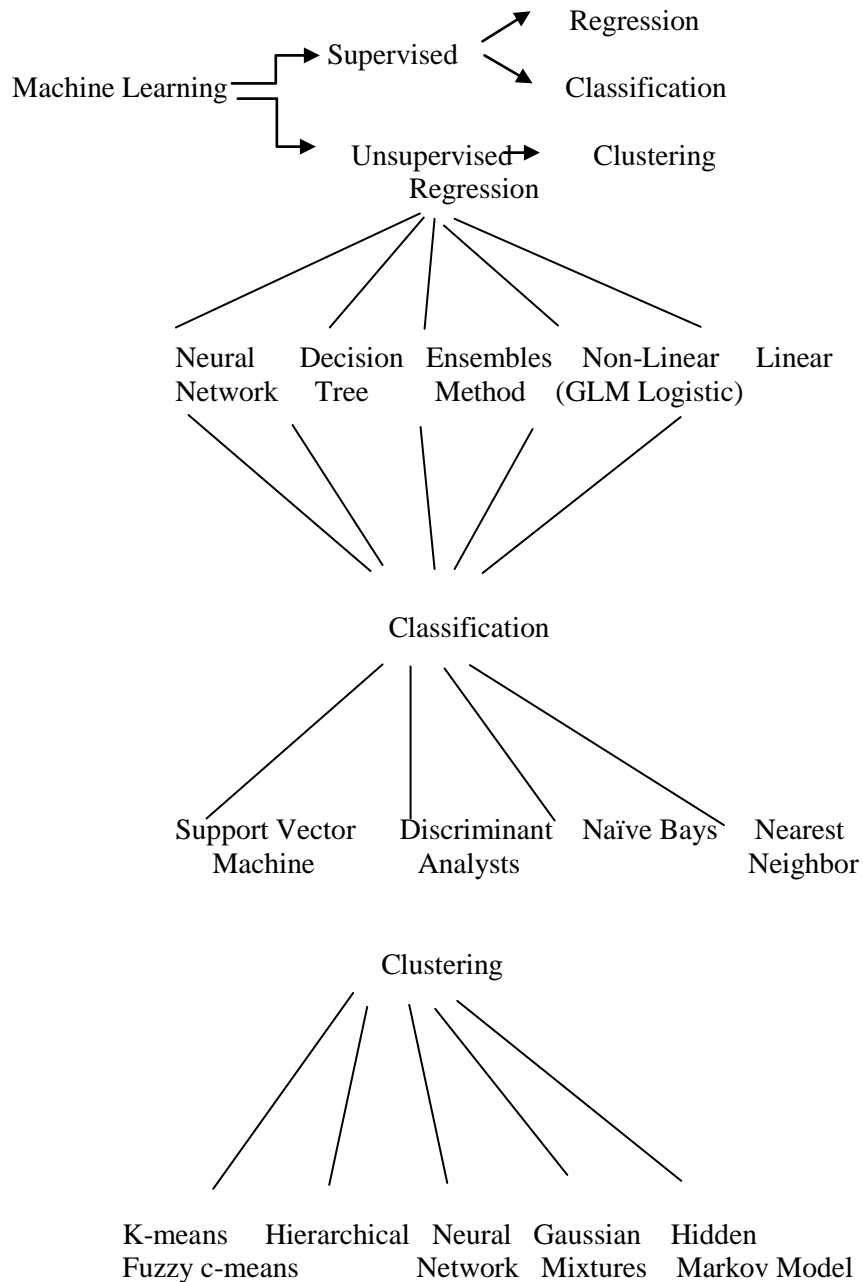
Learning can be implemented by various methodology or techniques depending upon the requirement. Basically learning can be categorized in two types -

- 1-supervised learning
- 2-unsupervised learning

In supervised learning two techniques are used

- 1-Regression
- 2-Classification

While unsupervised learning is implemented by clustering algorithms.



3. Applications

According to E. Don Box *et al.* [4] neural networks model can be used for demand forecasting in deregulated environment. Neural networks are designed and trained on the basis of total requests of the gatherings of studied clients of various classes.

The most every now and again experienced basic leadership undertakings of human action is classification. Classification issue happens when an article should be appointed into a predefined class in view of various observed attributes related to that object. Many problems in business, science, industry, and medicine can be treated as classification problems. Examples include bankruptcy prediction, medical diagnosis, credit scoring, quality control, speech recognition and handwritten character recognition [5].

Sankar K. Pal *et al.* [6] described the application of web mining in soft computing framework. There are various applications of Neural Network and genetic algorithms in web mining.

Soft computing paradigm like fuzzy sets (FS), artificial neural networks (ANN) and support vector machines (SVMs) is used in Bioinformatics [7].

The research community had work to discover IP traffic classification methods that don't fully depend on 'surely understood' TCP or UDP port numbers, or translating the substance of packet payloads. New work is developing on the utilization of statistical traffic attributes to bolster in the identification and classification process [8].

The capacity to anticipate or find out business disappointments is urgent for money related organizations, as wrong choices can have direct budgetary results. There are the two noteworthy examination issues in the accounting and finance area are Bankruptcy prediction and credit scoring. Various models have been created in the writing, to anticipate whether borrowers are in risk of bankruptcy and whether they should be considered a good or bad credit risk. Machine-learning techniques, such as neural networks and decision trees have been applied widely as a tools for bankruptcy prediction and credit score modelling Since the 1990s [9].

Learning methods that have been applied to CRs classifying them under supervised and unsupervised learning. Some of the most important, and commonly used, learning algorithms was provided along with their advantages and disadvantages are discussed in this literature [10].

Representation learning is a way of pattern analysis and design of intelligent machines [11].

Big data is the very challenging deep learning approach and used for digital object identification [12].

A Machine Learning Based Automotive Forensic Analysis Applications for Mobiles Using Data Mining *i.e.* automotive legal investigation of mobile applications in view of the produced multifaceted information is directed by machine learning is proposed by the framework [13].

Doubly Fed Induction Generator is controlled by Hybrid Artificial Neural Network. With the expanding utilization of wind force era, it is required to prompt the dynamic performance analysis of Doubly Fed Induction Generator under different working conditions [14].

Time Series Prediction Using Radial Basis Function Neural Network. It is a methodology utilized for anticipating day by day system movement utilizing artificial neural networks (ANN), to be specific radial basis function neural network (RBFNN) method [15].

Deep colorization is the latest example of deep machine learning application. Colours in black and white images can be added by deep machine learning. Process of image colorization uses very large convolution neural network (CNN) and supervised layers for recreating the image with colours [16].

There are also many applications of deep machine learning in modern era like Addition of Sounds to Silent Movies, Automatic Machine Translation, Classification of object in Photographs, Automatic Handwriting Generation, Character Text Generation, Image Caption Generation, Automatic Game Playing.

Google's Deep Mind project shows the success of deep machine learning technique. DeepMind claims that their framework is not pre-customized: it gains as a matter of fact, utilizing just crude pixels as information input. Technically it uses deep learning on a convolutional neural network, with a novel form of Q-learning, a form of model-free reinforcement learning [17][18].

4. Results and Analysis

In this paper we have discussed various machine learning techniques and their implementation *i.e.* supervised, unsupervised, discriminative, generative and deep machine learning and their implementation method like regression, classification and clustering *etc.* Regression problem can be solved by neural network, decision tree *etc.* while classification problems can be solved by neural network, decision trees, support vector machine, naïve bay's and nearest neighbor *etc.* Clustering problems can be solved by k-means, hidden markove models *etc.*

Now it is clear that different techniques are used for solving the problems depending upon the category or type of the problem whether it is classification problem or regression or clustering. In this survey paper we find out that deep machine learning which is extended version of supervised machine learning is used to solve the complex classification problem. Deep machine learning technique is implemented in two ways

1-Convolutional Neural Network

2-Deep Belief Neural Network

It is also found that accuracy of the result of deep learning methods (Convolution neural network and Deep Belief network) are better than traditional artificial neural network like perceptron and feed forward back propagation neural network. Deep machine learning provides solutions to various problems those were not solved by traditional methods like image colorization, it was done by human effort manually.

When we discuss large deep neural network (LDNNs), we are discussing 10-20 layer neural systems (since this is the network that can be trained with today's algorithms). we can give a couple of methods for taking a gander at LDNNs that will light up the reason they can do and they do.

- Customary statistical models learn basic examples or groups. Interestingly, LDNNs learn calculation, though a hugely parallel calculation with an unassuming number of steps. Without a doubt, this is the key distinction amongst LDNNs and other statistical models.

- To expand further: it is outstanding that any algorithms can be implemented by a suitable profound circuit (with a layer for each time step of the algorithm's execution). Besides, more profound the circuit, the more costly are the algorithms that can be implemented by the circuit (regarding runtime).

Surprisingly, neural systems are very proficient than Boolean circuits. By more effective, we imply that a genuinely shallow DNN can tackle issues that require numerous more layers of Boolean circuits. For a particular case, consider the profoundly amazing reality that a DNN with 2 hidden layers and an unobtrusive number of units can sort N-bit numbers! It was found that the outcome stunning when actualized a little neural system and was prepared to sort 10 6-bit numbers, which was anything but difficult to do amazingly. It is impossible to sort N N-bit numbers with a Boolean circuit that has two hidden layers and that are not gigantic.

The reason of DNNs are more productive than Boolean circuits is on account of neurons play out a threshold operation, which is impossible with a little Boolean circuit.

Therefore, it is clear that Deep Neural Networks are better than Boolean circuits also.

5. Conclusion

In this paper a thorough discussion about machine learning methods and their implementation and analytical discussion on deep machine learning and deep neural network has been mentioned. It is clearly shown that different methods uses different algorithm for implementation. It is also concluded that Neural Network and Support

vector machine is most popular techniques to implement the machine learning processes. Deep learning is extended version of supervised learning. It is finally concluded that Convolution neural network (CNN) and Deep Belief network (DBN) are two powerful techniques, which may be used to solve various complex problems using deep learning. Deep learning platforms can also be benefited from engineered features while learning more complex representations, which engineered systems typically lack. It is also described that Deep Neural Network (DNN) can solve the problems that are difficult to solve by Boolean circuits. It is copiously clear that headways made concerning growing deep machine learning frameworks will without a doubt shape what's to come in machine learning and artificial intelligence systems in general.

References

- [1] L. Deng and X. Li, "Machine Learning Paradigms for Speech Recognition: An Overview", *IEEE Trans On Audio, Speech, And Language Processing*, vol. 21, no. 5, (2013).
- [2] <http://cs231n.github.io/convolutional-networks/#overview>.
- [3] <http://image.slidesharecdn.com/deep-belief-nets1166/95/deep-belief-nets-3-728.jpg?cb=1272282825>.
- [4] C. Wiktor, B. E. Don, W.-J. Lee, M.-S., K. Chen Paul and O. Peter Van, "Neural-Network-Based Demand Forecasting in Deregulated Environment", *IEEE Trans. On Industry Applications*, vol. 36, no. 3, (2000).
- [5] G. P. Zhang, "Neural Networks for Classification: A Survey", *IEEE Trans. On Systems, Man, And Cybernetics—Part C: Applications and reviews*, vol. 30, no. 4, (2000).
- [6] S. K. Pal, V. Talwar and P. Mitra, "Web Mining in Soft Computing Framework: Relevance, State of the Art and Future Directions", *IEEE Trans. On Neural Networks*, vol. 13, no. 5, (2002).
- [7] S. Mitra and Y. Hayashi, "Bioinformatics with Soft Computing", *IEEE Trans. On System, Man and Cybernetics—Part C: Application and Reviews*, vol. 36, no. 5, (2006).
- [8] T. T.T. Nguyen and G. Armitage, "A Survey of Techniques for Internet Traffic Classification using Machine Learning", *IEEE Communications Surveys & Tutorials*, vol. 10, no. 4, (2008).
- [9] W.-Y. Lin, Y.-H. Hu and C.-F. Tsai, "Machine Learning in Financial Crisis Prediction: A Survey", *IEEE Trans. On System, Man and Cybernetics—Part C: Application and Reviews*, vol. 42, no. 4, (2012).
- [10] M. Bkassiny, Y. Li and S.K. Jayaweera, "A Survey on Machine-Learning Techniques in Cognitive Radios", *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, (2013).
- [11] Y. Bengio, A. Courville and P. Vincent, "Representation Learning: A Review and New Perspectives", *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, (2013).
- [12] X.-W. Chen and X. Lin, "Big Data Deep Learning: Challenges and Perspectives", *Digital Object Identifier 10.1109/IEEE ACCESS (2014)*.
- [13] MD. Hussain Khan and G. Pradeepini, "Machine Learning Based Automotive Forensic Analysis for Mobile Applications Using Data Mining", *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 16, no. 2, (2015), pp. 350 ~ 354.
- [14] G. Venu Madhav and Y. P. Obulesu, "A New Hybrid Artificial Neural Network Based Control of Doubly Fed Induction Generator", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 5, no. 3, (2015).
- [15] G. Venu Madhav and Y. P. Obulesu, "Time Series Prediction Using Radial Basis Function Neural Network", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 5, no. 4, (2015), pp. 765~771.
- [16] R. Zhang, P. Isola and A. A. Efros, "Colorful Image Colorization", iv:1603.08511v4 [cs.CV] 2 Sep (2016).
- [17] V. Mnih and K. Kavukcuoglu "Human-level control through deep reinforcement learning", *Nature*, vol. 518, (2015), pp. 529–33.
- [18] K. Volodymyr, S. Koray, G. David, A. Alex, W. Ioannis, R. Daan and Martin, "Playing Atari with Deep Reinforcement Learning", Retrieved 12 October, (2014).

Authors



Chandrahas Mishra, he completed M.Tech. Hons (2016) in Computer Science and Engineering from Kamla Nehru Institute of Technology, Sultanpur affiliated to Dr.A.P.J Kalam Technical University Lucknow (U.P). He received B.Tech. (2012) in Computer Science & Engineering from Accurate Institute of Management & Technology Greater Noida (U.P.). His research area is Machine Learning and Neural Networks, Artificial Intelligence, Data Mining and Image Processing *etc.*



Dharmendra Lal Gupta, he is currently working as an Associate Professor in the Department of Computer Science & Engineering at KNIT, Sultanpur (U.P.) India. He received B.Tech. (1999) from Kamla Nehru Institute of Technology (KNIT) Sultanpur (U.P.), in Computer Science & Engineering, M.Tech. Hons (2003) in Digital Electronics and Systems from Kamla Nehru Institute of Technology (KNIT) Sultanpur (U.P.) India. He has been a member of IEEE Computer Society. He has published about 28 papers in International/National Journals/workshops/conferences and seminars. His research interests are Software Quality Engineering, Software Engineering, Cryptography and Network Security.

