Smarter Artificial Intelligence with Deep Learning

Dr.V.V.Narendra Kumar[#1], T.Satish Kumar[*2]
 # Professor CSE, Department of Computer Science & Engineering,
**Asst.Professor, Department of Electronics & Communication Engineering,*
St.Mary's Engineering College, Hyderabad, India

Abstract

The unpredictable growth in large-scale computing capabilities, availability of large datasets, and advancements in learning techniques etc. made it necessary for Deep Learning. The rapid growth in the mentioned areasresulted in varied deep learning frameworks. But there are several inefficiencies in these frameworks in user and developer point of view.Moreover, adopting useful techniques across frameworks in performing learning tasks and optimizingperformance has become very essential.

Deep learning (DL) is a set of diversified approacheswhere machine learning can be innovative and to helping computers to usebig data i.e., huge amounts of data which is in the form of text, images and sound. Deep networks can be trained with vast amounts of data using deep learning algorithms. High level abstractions in data can be modelled using deep learning based on a set of algorithms. It is a new research area where Machine Learning can be drawn nearer to Artificial Intelligence. DL is used in various fields for achieving multiple levels of abstraction like sound, text, images feature extraction etc. Deep Learning is used by popular search engines like Google in its voice and image recognition algorithms, and by Netflix and e-commerce websites like Amazon, to decide what consumer wants to buy next, and even by researchers at MIT in predicting the future. Hence Deep Learning gained much significance in recent days. Many Universities started various courses in Deep Learning which indicates the importance of Deep Learning in the academic world.

Keywords - *Deep learning, deep machine learning, Supervised Learning, Artificial Intelligence, Artificial Neural Networks*

I. INTRODUCTION

Deep Learning [1-4] is a subset of machine learning, and machine learning is a subset of Artificial Intelligence (Fig. 1). Deep learning, Machine Learning and Artificial Intelligence is similar to a set of Russian dolls nested within each other called "**matryoshka doll**", beginning with the smallest and working out. **Machine learning** is one application of **Artificial Intelligence** (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves. **Artificial Intelligence** is the science and engineering of making intelligent machines. Artificial Intelligence is a branch of computer science dealing with the simulation of intelligent behaviour in computers.

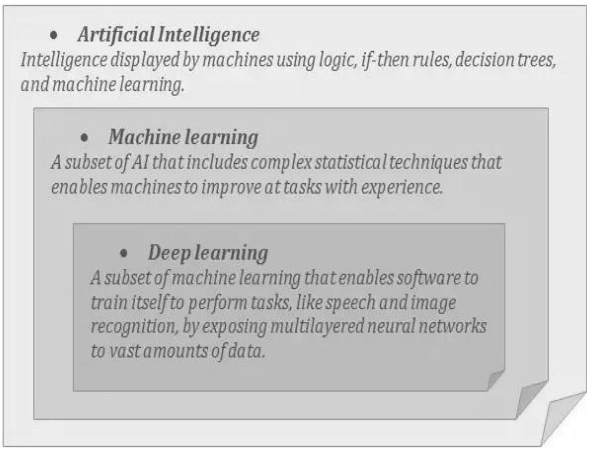


Fig.1. Artificial Intelligence, Machine Learning and Deep Learning

II. ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND DEEP LEARNING

Artificial Intelligence (AI)[5]is theability of a machine to imitate or simulate intelligent human behaviour. Using AI, a computer will be able to perform various tasks that requires human intelligence namely, visual perception, speech recognition, decision-making, and translation between languages etc.All such tasks basically require several "if-then" rules.AI basically combines huge amounts of data with rapid, iterative processing. It uses intelligent algorithms through which the software learns automatically from patterns or features in the data.Computers with Artificial Intelligence are designed for various activities such as, Speech Recognition, Learning, Planning, Problem Solving etc.

Through Turing test (Allan Turing), the intelligence of a computer can be determined (Fig. 2). In the Turing test, a computer can be said to be intelligent if it can achieve human-level performance in all cognitive tasks(i.e., the mental action or process of acquiring knowledge and understanding through thought, experience, and the senses.,) sufficient to fool an interrogator. In order to be artificially intelligent and pass the Turing test the computer should posses the following:

- **Natural Language Processing** (NLP) to enable it to communicate successfully in English (or some other human language).
- **Knowledge Representation** (KR) to store information provided before or during the interrogation.
- **Automated Reasoning** (AR) to use the stored information to answer questions and to draw new conclusions.

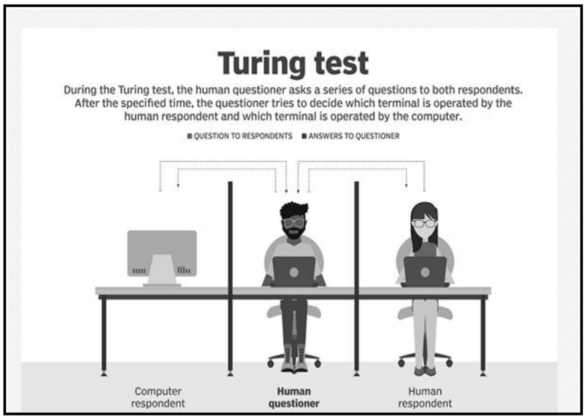


Fig.2.The process of a Turing test

All machine learning can be said as AI, but the vice versa may not be true i.e., not all AI is as machine learning. Hence **Machine learning (ML**)[3] is a subset of AI.For instance, symbolic logic – rules engines, expert systems and knowledge graphs –all are AI but all these arenot machine learning.Afeature that separates machine *learning* from the knowledge graphs and expert systems is its ability to change itself when exposed to more data. Machine learning is dynamic and does not require human intervention to make certain changes.

ING

As per Tom Mitchell, "A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E."

In machine learning, a certain dimension of a machine is optimized usinga variety of algorithms. This is called as the "learning" part of a machine. Hence computers try to minimize error or maximize the likelihood of their predictions being true through machine learning. This is called as an error function, a loss function, or an objective function. From the objective function the machine learning algorithm one can get the general idea of its value.

Error can be minimized by building a framework which multiplies inputs in order to make guesses as to the inputs nature. The product of the inputs and the algorithm is nothing but various outputs or guesses. Generally the initial guesses can be erroneous, but if you have empirical evidencepertaining to the input, you can measure the deviation of yourpresumptions by contrasting them with the facts, and then use that error to modify your algorithm. This is what happens in neural networks. In neural networks error parameters are measured and modified continuously until they can't achieve any less error.These are called optimization algorithms where they minimize their error by several continuous guesses.

Machine learning is a process of data analysis that automates analytical model building. A class of artificial intelligence (AI) that enables software applications to become more accurate in forecasting outcomes without being specially programmed is nothing but Machine learning.The main idea of machine learning is to create algorithms that can receive input data and use statistical analysis to predict an output value within an acceptable range.

The processes that are involved in machine learning are similar to data mining and predictive modelling where they require searching through data to look for patterns and rectifying program actions suitably. Customers who shop online like in Amazon, Flipkart etc. are familiar with machine learning from the products being offeredrelated to their purchase. Most industries realize advantages of machine learning technology by collecting insights from huge data they collect in real time and companies can work more efficiently or gain competitive edge.

Google uses Machine Learning very widely. For example, when Google Music or YouTube suggests what other title or video you might be interested in. If you use Google Inbox for your emails, Inbox provides "smart replies" (Fig.4.) which devises possible replies you might be sending back. The new Google Translate premium used a neural machine translation system to make native-sounding translations of text.

Fig.3. Google Machine learning apps

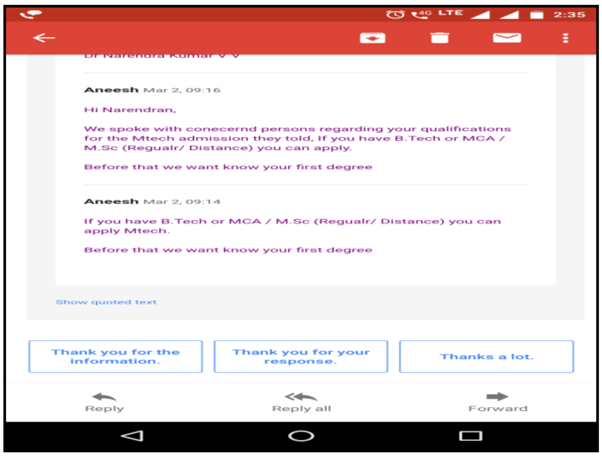


Fig.4. Smart replies in Gmail

Coming to **deep learning,** it is a subset of machine learning [6]. Other terms for deep learning aredeep artificial neural networks or also deep reinforcement learning.

Deep artificial neural networks [6] are a set of algorithms that have set new records in accuracy for many important problems, such as image recognition, sound recognition, recommender systems, etc. For example, deep learning is part of DeepMind's well-known AlphaGo algorithm, which beat the former world champion Lee Sedol at Go in early 2016, and the current world champion KeJie in early 2017.

GET STARTED WITH DEEP LEARNING

Deep is a technical term. It refers to the number of layers in a neural network. A shallow network has one so-called *hidden layer*, and a deep network has more than one. Multiple hidden layers allow deep neural networks to learn features of the data in a so-called feature hierarchy, because simple features (e.g. two pixels) recombine from one layer to the next, to form more complex features (e.g. a line). Nets with many layers pass input data (features) through more mathematical operations than nets with few layers, and are therefore more computationally intensive to train. Computational intensivity is one of the hallmarks of deep learning, and it is one reason why GPUs are in demand to train deep-learning models.

One can apply the same definition to deep learning that Arthur Samuel did to machine learning – a "field of study that gives computers the ability to learn without being explicitly programmed" – while adding that it tends to result in higher accuracy, require more hardware or training time, and perform exceptionally well on machine perception tasks that involved unstructured data such as blobs of pixels or text.

The difference between AI and Deep Learning, which is currently at the border of our understanding of AI, and Machine Learning, is the concept of *generalization*.

Machine Learning is about finding patterns, dependencies and invariants in (big quantities of) data. Before initiation of Artificial Neural Networks (ANN) this is where the focus was, in finding and understanding those hitherto unknown relationships.ANNs raised the bar of machine learning by displaying generalizations in ways that were often not understood how deep networks did it. Recent advances in Deep Learning raised it further still by using much deeper networks with much more data, being deployed at tremendous scale by leading technology companies such as Google, Face book and others.

We still do not really fully understand how these deep learning networks work, even though progress is being made constantly in that direction too. The key part of recent deep learning breakthroughs is accepting the results without fully understanding how they do it.

A. Examples of Deep Learning

Deep learning applications [7] are used in industries from automated driving to medical devices.

1. Automated Driving*:* Automotive researchers are using deep learning to automatically detect objects such as stop signs and traffic lights. In addition, deep learning is used to detect pedestrians, which helps decrease accidents.

2. Aerospace and Defence: Deep learning is used to identify objects from satellites that locate areas of interest, and identify safe or unsafe zones for troops.

3. Medical Research*:* Cancer researchers are using deep learning to automatically detect cancer cells. Teams at UCLA built an advanced microscope that yields a high-dimensional data set used to train a deep learning application to accurately identify cancer cells.

4. Industrial Automation*:* Deep learning is helping to improve worker safety around heavy machinery by automatically detecting when people or objects are within an unsafe distance of machines.

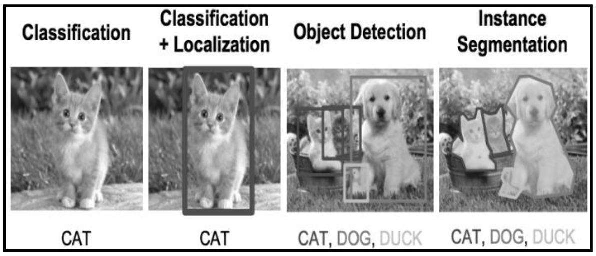
5. Electronics*:* Deep learning is being used in automated hearing and speech translation. For example, home assistance devices that respond to your voice and know your preferences are powered by deep learning applications.

B. Application of deep learning to computer vision-Types of Computer Vision Tasks

Computer Vision [8], as the name suggests is simply creating artificial models which can replicate the visual tasks performed by a human. This essentially means what we can see and what we perceive is a process which can be understood and implemented in an artificial system.

The main types of tasks that computer vision can be categorised in are as follows (Fig.5):

- **Object Recognition / classification** – In object recognition, you are given a raw image and your task is to identify which class the image belongs to.
- **Classification + Localisation** – If there is only one object in the image, and your task is to find the location of that object, a more specific term given to this problem is localization problem.
- **Object Detection** – In object detection, you task is to identify where in the image does the objects lies in. These objects might be of the same class or different class altogether.
- **Image Segmentation** – Image Segmentation is a bit sophisticated task, where the objective is to map each pixel to its rightful class.



III. DEEP LEARNING MODELS AND ARCHITECTURES

Deep learning algorithms consists of a diverse set of models in comparison to a single traditional machine learning algorithm. This is because of the flexibility that neural network provides when building a full-fledged end-to-end model.

Neural network can sometimes be compared with Lego blocks (Fig 6.), where you can build almost any simple to complex structure your imagination helps you to build.

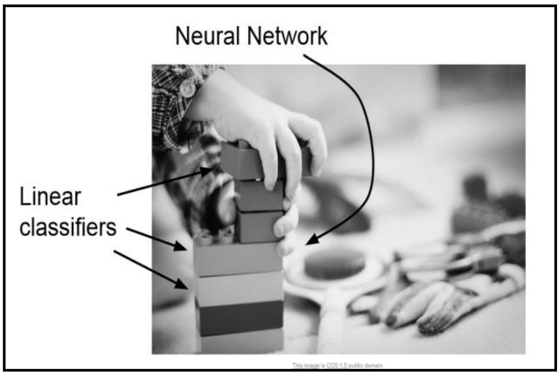


Fig.6. Lego blocks

We can define an advanced architecture as one that has a proven track record of being a successful model. This is mainly seen in challenges like ImageNet, where your task is to solve a problem, say image recognition, using the data given. Those who don't know what ImageNet is, it is the dataset which is provided in ILSVR (ImageNet Large Scale Visual Recognition) challenge.

Also as described in the below mentioned architectures, each of them has a nuance which sets them apart from the usual models; giving them an edge when they are used to solve a problem. These architectures also fall in the category of "deep" models, so they are likely to perform better than their shallow counterparts.

IV. DEEP LEARNING ARCHITECTURES

The most important Deep Learning Architectures are: AlexNet, VGG (Visual Graphics Group) Net, GoogleNet, ResNet (Residual Networks), ResNeXt, RCNN (Region Based CNN), YOLO (You Only Look Once), SqueezeNet, SegNet, SegNet. A brief description of GoogleNet is given below as an example:

A. GoogleNet

GoogleNet (or Inception Network) is a class of architecture designed by researchers at Google. GoogleNet was the winner of ImageNet 2014, where it proved to be a powerful model.

In this architecture, along with going deeper. GoogleNet has 22 layers, and almost 12x less parameter.Their idea was to make a model that also could be used on a smart-phone. The researchers also made a novel approach called the Inception module (Fig. 7).

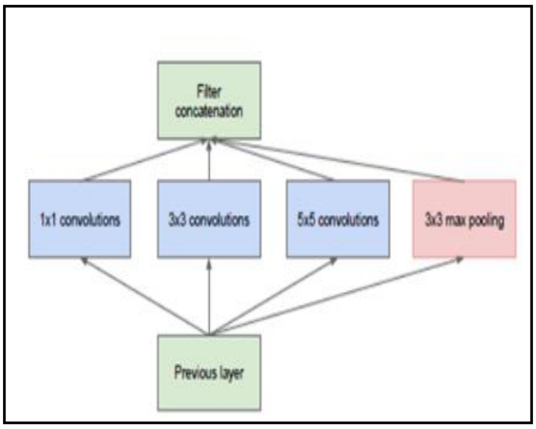


Fig.7. Inception Module, naive version

As seen above, it is a drastic change from the sequential architectures which we saw previously. In a single layer, multiple types of "feature extractors" (Fig. 8)are present. This indirectly helps the network perform better, as the network at training itself has many options to choose from when solving the task. It can either choose to convolve the input, or to pool it directly.

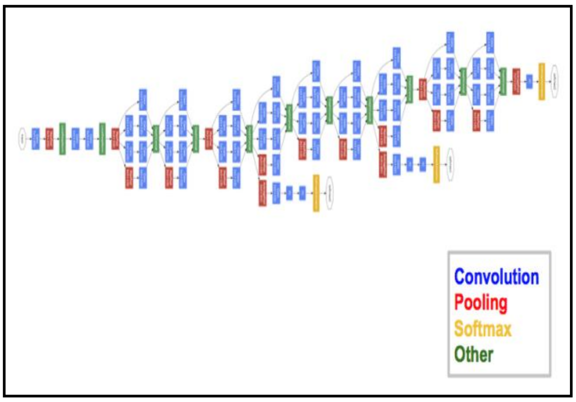


Fig.8. Feature Extractors in GoogleNet

The final architecture contains multiple of these inception modules stacked one over the other. Even the training is slightly different in GoogleNet, as most of the topmost layers have their own output layer. This nuance helps the model converge faster, as there is a joint training as well as parallel training for the layers itself.

The advantages of GoogleNet are:

- GoogleNet trains faster than VGG.
- Size of a pre-trained GoogleNet is comparatively smaller than VGG. A VGG model can have >500 MBs, whereas GoogleNet has a size of only 96 MB

GoogleNet does not have an immediate disadvantage per se, but further changes in the architecture are proposed, which make the model perform better. One such change is termed as an Xception Network, in which the limit of divergence of inception module (4 in GoogleNet as we saw in the image above) are increased. It can now theoretically be infinite (hence called extreme inception!)

B. Depth

The computations involved in producing an output from an input can be represented by a **flow graph**: a flow graph is a graph representing a computation, in which each node represents an elementary computation and a value (the result of the computation, applied to the values at the children of that node). Consider the set of computations allowed in each node and possible graph structures and this defines a family of functions. Input nodes have no children. Output nodes have no parents.

The flow graph for the expression **sin(a^2 +b/a)**could be represented by a graph with two input nodes **a** and **b**, one node for the division b/a taking **a** and **b** as input (i.e. as children), one node for the square (taking only **a** as input), one node for the addition (whose value would be **a^2+b** and taking as input the nodes **a^2** and **b/a**, and finally one output node computing the sinus, and with a single input coming from the addition node.

A particular property of such *flow graphs* is **depth**: the length of the longest path from an input to an output.

Traditional feed forward neural networks can be considered to have depth equal to the number of layers (i.e. the number of hidden layers plus 1, for the output layer). Support Vector Machines (SVMs) have depth 2 (one for the kernel outputs or for the feature space, and one for the linear combination producing the output).

C. Motivations for Deep Architectures

The main motivations for studying learning algorithms for deep architectures are the following:

- Insufficient depth can hurt
- The brain has a deep architecture
- Cognitive processes seem deep

1. Insufficient depth can hurt

Depth 2 is enough in many cases (e.g. logical gates, formal [threshold] neurons, sigmoid-neurons, Radial Basis Function [RBF] units like in SVMs) to represent any function with a given target accuracy. But this may come with a price: that the required number of nodes in the graph (i.e. computations, and also number of parameters, when we try to learn the function) may grow very large. Theoretical results showed that there exist function families for which in fact the required number of nodes may grow exponentially with the input size. This has been shown for logical gates, formal neurons, and RBF units. In the latter case Hastad has shown families of functions which can be efficiently (compactly) represented with **O(n)** nodes (for **n** inputs) when depth is d, but for which an exponential number ($O(2^n)$) of nodes is needed if depth is restricted to **d-1**.

One can see a deep architecture as a kind of factorization. Most randomly chosen functions can't be represented efficiently, whether with a deep or a shallow architecture. But many that can be represented efficiently with a deep architecture cannot be represented efficiently with a shallow one (see the polynomials example in the Bengio survey paper). The existence of a compact and deep representation indicates that some kind of structure exists in the underlying function to be represented. If there was no structure whatsoever, it would not be possible to generalize well.

The brain has a deep architecture. For example, the visual cortex is well-studied and shows a sequence of areas each of which contains a representation of the input, and signals flow from one to the next (there are also skip connections and at some level parallel paths, so the picture is more complex). Each level of this feature hierarchy represents the input at a different level of abstraction, with more abstract features further up in the hierarchy, defined in terms of the lower-level ones.

Note that representations in the brain are in between dense distributed and purely local: they are **sparse**: about 1% of neurons are active simultaneously in the brain. Given the huge number of neurons, this is still a very efficient (exponentially efficient) representation.

2. Cognitive processes seem deep

- Humans organize their ideas and concepts hierarchically.
- Humans first learn simpler concepts and then compose them to represent more abstract ones.
- Engineers break-up solutions into multiple levels of abstraction and processing

It would be nice to learn / discover these concepts (knowledge engineering failed because of poor introspection?). Introspection of linguistically expressible concepts also suggests a *sparse* representation: only a small fraction of all possible words/concepts are applicable to a particular input (say a visual scene).

V. BREAKTHROUGH IN LEARNING DEEP ARCHITECTURES

Before 2006, attempts at training deep architectures failed: training a deep supervised feed forward neural network tends to yield worse results (both in training and in test error) then shallow ones (with 1 or 2 hidden layers).

Three papers changed that in 2006, spearheaded by Hinton's revolutionary work on Deep Belief Networks (DBNs):

- Hinton, G. E., Osindero, S. and Teh, Y., A fast learning algorithm for deep belief nets Neural Computation 18:1527-1554, 2006
- YoshuaBengio, Pascal Lamblin, Dan Popovici and Hugo Larochelle, Greedy Layer-Wise Training of Deep Networks, in J. Platt et al. (Eds), Advances in Neural Information Processing Systems 19 (NIPS 2006), pp. 153-160, MIT Press, 2007
- Marc'AurelioRanzato, Christopher Poultney, Sumit Chopra and Yann LeCun Efficient Learning of Sparse Representations with an Energy-Based Model, in J. Platt et al. (Eds), Advances in Neural Information Processing Systems (NIPS 2006), MIT Press, 2007

The following key principles are found in all three papers:

- Unsupervised learning of representations is used to (pre-) train each layer.
- Unsupervised training of one layer at a time, on top of the previously trained ones. The representation learned at each level is the input for the next layer.
- Use supervised training to fine-tune all the layers (in addition to one or more additional layers that are dedicated to producing predictions).

The DBNs use RBMs for unsupervised learning of representation at each layer. The Bengio et al paper explores and compares RBMs and *auto-encoders* (neural network that predicts its input, through a bottleneck internal layer of representation). The Ranzato et al paper uses sparse auto-encoder (which is similar to *sparse coding*) in the context of a *convolutional* architecture. Auto-encoders and convolutional architectures will be covered later in the course.

VI. CONCLUSION

Deep Learning is used as a function evaluation component in a much larger search algorithm AlphaGo which employed Deep Learning in its value and policy evaluations. Google's Gmail auto-reply system used DL in combination with beam searching. More of these hybrid algorithms rather than new end-to-end trained DL systems are anticipated in future. End-to-end Deep Learning is a fascinating area of research, but for now hybrid systems are going to be more effective in application domains.Over the past few years, Deep Learning (DL) architectures and algorithms have made impressive advances in fields such as image recognition and speech processing.Their application to Natural Language Processing (NLP) was less impressive at first, but has now proven to make significant contributions, yielding state-of-the-art results for some common NLP tasks. Named entity recognition (NER), part of speech (POS) tagging or sentiment analysis is some of the problems where neural network models have outperformed traditional approaches. The progress in machine translation is perhaps the most remarkable among all.

Some anticipated Deep Learning trends in the future may be- hardware will accelerate doubling Moore's law, Convolution Networks (CNN) will Dominate, Designers will rely more on Meta-Learning, Reinforcement Learning will only become more creative, Adversarial and Cooperative Learning will be ruler, Predictive Learning or Unsupervised Learning will not progress much, More Applications will use Deep Learning as a component, and Design Patterns will be increasingly Adopted, Engineering will outpace Theory. In the coming days Deep Learning will play main role in most of the software and hardware components.

REFERENCES

[1]. Matthew N. O. Sadiku, Mahamadou Tembely, and Sarhan M. Musa, “Deep learning,” International Research Journal of Advanced Engineering and Science, Volume 2, Issue 1, pp. 77-78, 2017.
[2]. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." Nature 521.7553 (2015): 436-444.
[3]. L. Arnold et al., “An introduction to deep learning,” Proceedings of European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, pp. 477-488, April 2011.
[4]. W. Wang et al., “Deep learning at scale and at ease,” ACM Transactions on Computer Communication Applications, vol. 12, no. 4s, pp. 69:1- 69:25, Nov. 2016.
[5]. Arel, I., Rose, D. C., &Karnowski, T. P. (2010). Deep machine learning-a new frontier in artificial intelligence research [research frontier].Computational Intelligence Magazine, IEEE, 5(4), 13-18.
[6]. Schmidhuber, Jurgen. "Deep learning in neural networks: An overview." Neural networks 61 (2015): 85-117.
[7]. SSSN Usha Devi N,Dr R. Mohan,Dr P.KiranSree, “An Extensive Survey on Deep Learning Applications”, International Journal of Innovations & Advancement in Computer Science IJIACS ISSN 2347 – 8616 Volume 6, Issue 2 February 2017, 81-84.
[8]. Wan J, Wang D, Hoi SCH, Wu P, Zhu J, Zhang Y, et al. Deep learning for content-based image retrieval: a comprehensive study. In: ACM MM; 2014. p. 157–66.