International Journal of Latest Trends in Engineering and Technology
Vol.(10)Issue(1), pp.379-384
DOI: http://dx.doi.org/10.21172/1.101.69
e-ISSN:2278-621X

DEEP LEARNING: OVERVIEW,ARCHITECTURE, FRAMEWORK & APPLICATIONS

Mamta Arora[1], Sanjeev Dhawan[2], Kulvinder Singh[3]

Abstract- Deep learning is a branch of artificial intelligence that is used to solve the complicated functions which are represented by higher level of abstraction such as computer vision, natural language processing, driverless cars, etc. The remarkable results have shown the rapid increase in usage of deep learning techniques in challenging and complex area. The varieties of frameworks have been developed by different organizations like Google, Facebook, and Intel for assisting the programmer in modeling the solution of the problem. The family of deep learning architectures is broadly classified into fully connected network, Convolutional network and recurrent neural network. Each of the deep learning neural network architecture is having distinguished feature.
Keywords – Machine Learning, Neural networks, deep Learning, architecture

1. INTRODUCTION

Machine learning (ML) play a vital role in designing many applications that can serve better to society: from recommendation on e-commerce websites to content filtering on social network , giving suggestion for making friends on Facebook by creating network of friends of friends. Not only for web applications it is also embedded in consumer products like medical wristbands [1] to identify patient, transcribe speech into text, camera and smart phones. Progressively, applications in these areas are using techniques of deep learning. ML techniques includes conventional algorithms like Decision Tree, Classification and Regression Tree, Naive Bayes etc. whereas Deep learning (DL) techniques includes Auto encoder neural network, Deep Stack network(DSN),Deep Belief Network(DBN) Long Short term Memory (LSTM), Extreme Learning Model (ELM), Generative Adversarial Networks (GANs) etc. The distinguish facet of deep learning algorithms is leaning of features without human intervention. Although ML progressed rapidly in recent years but DL algorithms

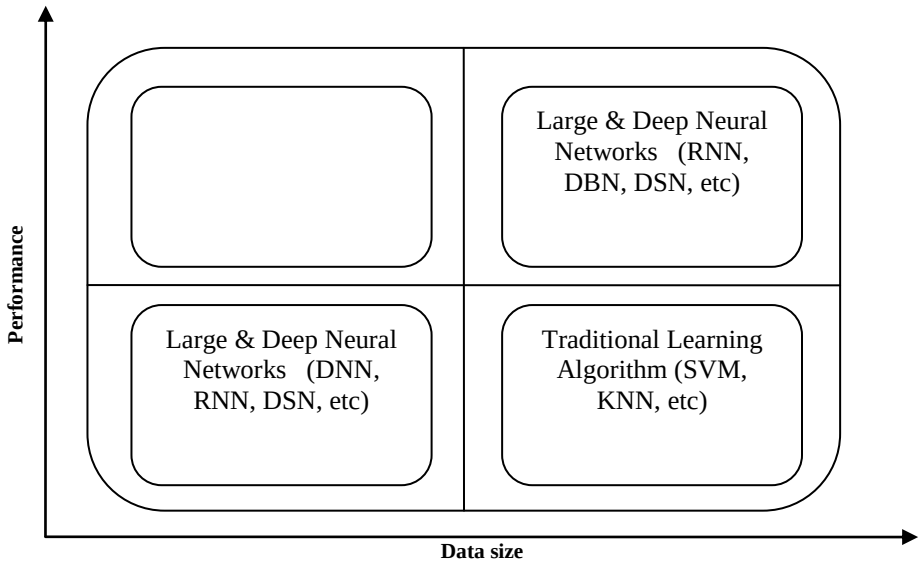

Figure 1 Performance of algorithms with respect to data

[1] Ph.D. Research Scholar, Department of Computer Science & Engineering University Institute of Engineering & Technology (U.I.E.T), Kurukshetra University, Kurukshetra,Haryana,India
[2] Faculty of Computer Science & Engineering,Department of Computer Science & Engineering,University Institute of Engineering & Technology (U.I.E.T), Kurukshetra University, Kurukshetra, Haryana,India
[3] Faculty of Computer Science & Engineering,Department of Computer Science & Engineering,University Institute of Engineering & Technology (U.I.E.T), Kurukshetra University, Kurukshetra, Haryana,India

outperforms better where we have less or no domain expertise knowledge. For example to identify pattern in an image machine learning algorithm require careful selection of feature and the tuning of features which is time consuming job. On the other hand deep learning algorithms fed with raw data .i.e. array of pixels and will learn feature automatically.

In commencement of big data, traditional machine learning algorithms were highly efficient and effective. This is because traditional algorithms such as random forests, support vector machines, and even logistic regressions improve very quickly when you add data. However, their performance plateaus after hitting millions of examples and the margin of improvement in performance related to data size become minimal. To overcome this limitation deep learning came into existence that can give better performance with the bigger size of data. The figure 1 depicts the relationship of data size and performance in respect to traditional ML algorithms and DL algorithms. The fourth quadrant shows improved performance with more data size by applying the DL algorithms.

The figure 2 depicts the current trends of DL, AI and ML for the last 5 years in healthcare domain from research and technology perspective. As with the advanced tools and techniques used in healthcare sector we are getting more and more electronic and digital data. So there is need of paradigm shift from ML to DL for better diagnosis and predictions. This paper is organized three main sections. The second section gives the overview of architecture of deep learning networks which gives the overview of each model along with their applications. The third section consolidates the commonly used open source frameworks that help in implementation part. The fourth section lists the various remarkable results of deep neural networks.

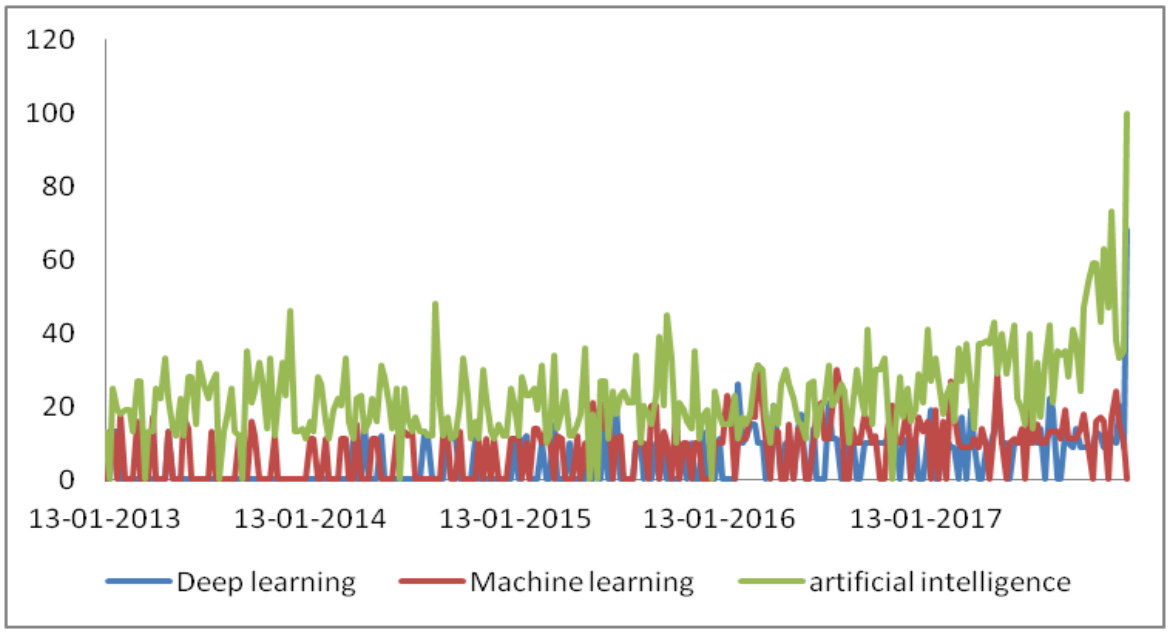

Figure 2 Trends: Deep Learning vs. Machine Learning vs. Artificial Intelligence [7]

2. ARCHITECTURE: NEURAL NETWORK AND DEEP LEARNING

The Artificial neural networks (ANNs) inspired by human biological nervous system. The main computational unit of human brain is a neuron. One of the most former neural networks which was based on human brain system is Perceptron [2]. It is having a very simple architecture which is consist of two layers: input layer and output layer. The units at input layer are directly connected to the units of output layer. It is best suited for classifying linear separable patterns. For solving the complex patterns, it uses layered architecture which is consisting of input layer, hidden layers and output layer. The fully connected neural network is depicted in figure 3 where each input node is connected with every other node in hidden layer with certain weight Wij and which in-turn is connected with every other node in output layer with some weight Wjk. The hidden layers perform constructive computations intermediately before passing the input values to the output layer [3].

The addition of more hidden layers will better able to serve in finding the complex relationships. The Deep learning implies Deep Neural Network that is characterized by the depth factor. The depth factor of deep neural network indicates the number of hidden layer that can range from two to thousands.

A wide variety of deep neural network does exists presently that can be selected depending upon the application and problem domain. Some of them which are widely used for image detection, natural language processing ,scene labelling are described in detail along with the architecture.

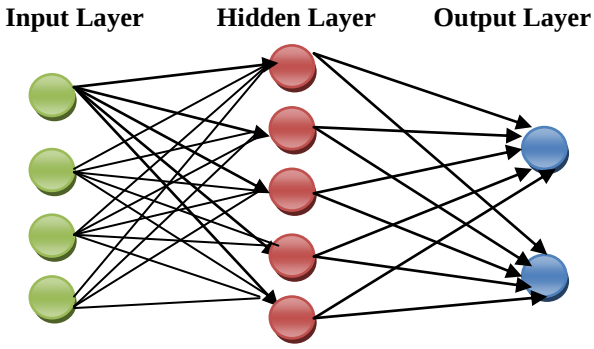

Figure3 Neural Network Architecture

2.1. Convolutional Neural Networks (CNNs/ConvNet)

Convolutional network is very much similar to the regular layered neural network which is having input layer, one or more hidden layer and output layer. The property that makes it different from regular NN is that it makes the explicit assumption that input are images, which allows to encode certain properties into architecture. It uses Convolutional filters that transform the 2D input data into 3D data. The architecture is shown in Figure 4. ConvNet architecture is consisting of three major layers: Convolution Layer, Pooling layer and fully connected layer. Each Layer accepts an input 3D volume and transforms it to an output 3D volume through a differentiable function. This model is having fast learning and good performance. The classification accuracy of this model is directly related to size of labelled data. The CNN is having benchmarked application in domain of Image recognition, video analysis, natural language processing.

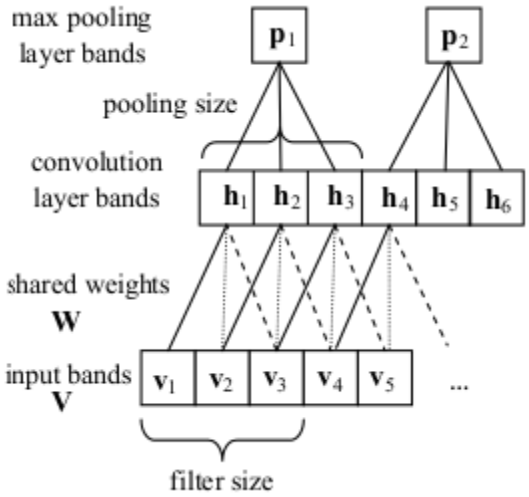

Figure3 Convolutional Neural Network [8]

2.2. Recurrent Neural Networks(RNNs)

The RNN is the foundation of all the deep learning architecture. The connection among the neurons makes it different from the multilayer feed forward network. As in RNN the neurons can be connected with the neuron of the other layers as well as with the neurons of the same layer for providing the feedback. The feedback connection supports RNNs in maintaining memory of past inputs and thus model problems in time. The architecture of RNN is shown in the Figure 5. This network is having many variations which include Long short term memory networks(LSTM), Bidirectional Long Short-Term Memory (BLSTM) neural networks ,Multi dimensional LSTM (MDLSTM) and hierarchical long-short term memory recurrent network (HLSTM).These networks performs better with the availability of big data. The widely used application of RNN is in the area of Speech recognition, handwriting recognition.

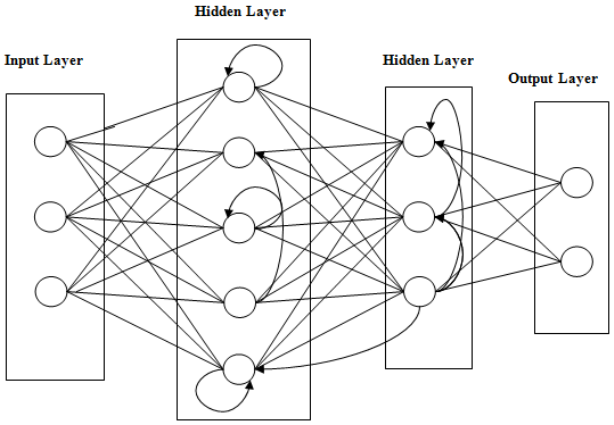

Figure4 Recurrent neural network

2.3. Deep Belief Networks(DBNs)

The DBN is similar to multilayer network having input layer, many hidden layers and output layer. Each pairs of layer in DNN is connected and works as restricted Boltzmann machine (RBM) as shown in figure 6. Thus DNN can be thought as stack of RBMs. The training in this network is carried out in two phases: pre training and fine tuning. In Pre-training phase each RBM is trained to reconstruct its input and this input will be fed to next RBM. This process will continue until each layer is pertained. Once the pertaining phase is over the network will undergo with fine tuning where the output layer is used to label the network. The major applications of DBN includes Image recognition, Scene labeling, information retrieval, natural language processing, failure prediction.

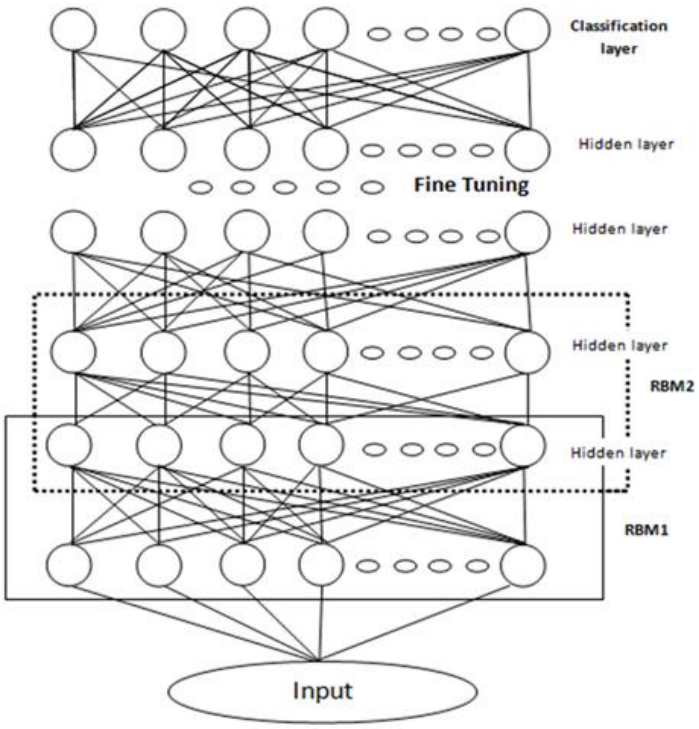

Figure5 Deep Belief Network

2.4. Deep Stacking Network (DSNs)

One of the major problems of DNN is that its computational processing cost increase exponentially with every new hidden layer. To overcome the limitation of DNN the DSN came with the idea of modularity. The architecture of DSN is divided into sub networks, where each sub-network is contains only one hidden layer as shown in the figure 7. The sub-net is also termed as module. There are two subnets one is shown in blue and other is shown in red color. The W and U represent the weight matrices which are connected through hidden layer. The modular architecture of DSN provides higher parallelism as output of any subnet can be copied to any other subnet of higher level. The foremost application areas of DSNs are Information retrieval, continuous speech recognition.

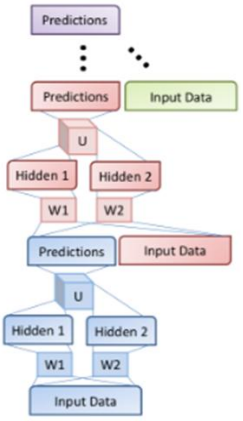

Figure6 Deep Stack Network

3. DEEP LEARNING FRAMEWORKS

A framework is environment that is built by system software to give platform to programmer for developing and deploying their applications. Deep Learning Frameworks enable the programmer to built and test their deep learning based applications. There are wide variety open source deep learning frameworks available in the market such as TensorFlow, Keras, Caffe, Torch, pyTorch etc. Some of the most commonly used frameworks are listed in Table1. Each row in the table is correspond to one open source framework which are attributed based on Developer's group, supported language and best suitable application.

Table I gives the clear view that based on the language known we can select framework. For example if we are comfortable with Java then we can choose DeepLearning4j (DL4j) for developing application specific software. All of the listed framework support python language which is according to TIOBE index [4] is fifth most powerful language. If we wish to design the application for the big data that can also run on distributed mode then we can use BigDL framework.

4. APPLICATIONS

Deep learning is making remarkable advances in giving the solutions to the problems that have shown the best attempts of the artificial intelligence community for many years [5]. Some of the benchmarked applications are listed in table 2. The performance column contains entry for accuracy and top-5 error. Top-5 error indicates prediction error out of top 5 classes.

5. CONCLUSION

In this paper, an attempt has been made to present the state of art of deep learning and its success in various image classification, scene labeling, and human body pose recognition. The commonly used deep learning networks are explained with their pros and cons. Each of the architecture is performing best based on the underlying distinguished feature. Deep learning networks can be easily incorporated in the solution by using the available open source frameworks. The commonly used frameworks are well explained in this paper.

Table 1: Deep learning open source frameworks

Name	Develop by	Interface	Major Application Domains
TensorFlow	Google	Python, C,C++,Java, Go, R	Voice & Sound Recognition, Video Detection, Time Series, Image Recognition, Text Summarization
Keras	Google	Python, R	Prediction, Feature extraction, and Fine-tuning
Caffe2	Facebook	Python, MatLab	Image classification and image segmentation ,RNNs
Torch	Swiss Institute EPFL	Lua, C, OpenCL	Used for implementing complex Neural Network topologies
PyTorch	Facebook	Python	Natural language processing
DeepLearning4j(DL4j)	ML Group, San Francisco	Java, Scala, Python,Kothlin and Clojure	Security applications, Recommender System, parallel and distributed application
MXNet	Apache	C++, Python, Julia, Matlab, Ja vascript, Go, R, Scala, Perl, W olfram Language	Used for Scalable Enterprise applications, supports flexible programming model with multiple languages

Microsoft Cognitive Toolkit	Microsoft Research	Python, C++,Command Line, Brain script	Image, handwriting and speech recognition problems
Deeplearn.Js	Google	Java Script,	For building distributed applications on Hadoop and spark framework, running application on browser
BigDL	Intel	Scala ,Python	Runs applications directly on the top of spark

Table 2: Summary of some deep learning based applications

Authors	Application Area	Data Source	Input Data	Deep Learning Method	Performance
Krizhevsky et al.[5]	Image Classification & Detection	ImageNet	1.2 million images	CNN	Top-5 Error=17.0%
Szegedy et al.[6]	Classification and detection	ImageNet	1.2 million images	GoogLeNet	Top-5 error 6.67%
Farabet et al.[7]	Scene Labeling	Standford Background dataset[15]	715 images outdoor scenes	Multiscale Feed Forward Supervised CNN	Per pixel accuracy 81.4%
Tompson et al.[8]	Human Body Pose estimation	FLIC [27]	5003 images	Hybrid: CNN with a Markov Random Field	-
Sainath et al.[9]	Large vocabulary continuous speech recognition (LVCSR)	News Channel	400 hours of English Broadcast News [15].	CNN	13% relative improvement over the Gaussian Mixture Model/Hidden Markov Model GMM/HMM[2]
Lei lei et al.	Cervical type identification	Kaggle test set	5278 labeled training images and 512 unlabelled test images	CNN	Accuracy 76%

6. REFERENCES

[1] L. V. Hoffmeister and G. M. S. S. d. Moura, "Use of identification wristbands among patients receiving inpatient," Revista Latino-Americana de Enfermagem(RLAE), vol. 23, no. 1, pp. 36-43, 2015.

[2] M. T. Jones, "index.html," IBM, 08 September 2017. [Online]. Available: https://www.ibm.com/developerworks/library/cc-machine-learning-deep-learning-architectures/index.html. [Accessed 17 January 2018].

[3] K. U. Rani, "Analysis of Heart Diseases using Neural network Approach," International Journal of Data Mining & Knowledge Management Process(IJDKP), vol. 1, no. 5, pp. 1-8, 2011.

[4] "Tiobe-index," [Online]. Available: https://www.tiobe.com/tiobe-index/. [Accessed 25 January 2018].

[5] Y. LeCun and Y. B. a. G. Hinton, "Deep Learning," nature, pp. 436-444, 2015.

[6] M. I. Razzak, S. Naz and A. Zaib, "Deep Learning for Medical Image Processing: Overview, Challenges and Future," in Classification in BioApps, Springer, Cham, 2018, pp. 323-350.

[7] "Trends," google, [Online]. Available: https://trends.google.com. [Accessed 7 January 2018].

[8] O. Abdel-Hamid, A.-r. Mohamed, H. Jiang, L. Deng, G. Penn and D. Yu, "Convolutional Neural Networks for Speech Recognition," IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, pp. 1-26, 2014.

[9] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going Deeper with Convolutions," Computer Vision and Pattern Recognition, pp. 1-9, 2014.

[10] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," COMMUNICATIONS OF THE ACM, pp. 84-90, 2017.

[11] C. Farabet, C. Couprie, L. Najman and Y. LeCun, "Learning hierarchical features for scene labelling," IEEE Trans. Pattern Analytics and Machine Intelligence, pp. 1915-1929, 2013.

[12] J. Tompson, A. Jain, Y. LeCun and C. Bregler, "Joint training of a convolutional network and a graphical model for human pose estimation," Advances in neural information processing systems, pp. 1799-1807, 2014.

[13] T. N. Sainath, A.-r. Mohamed, B. Kingsbury and B. Ramabhadran, "DEEP CONVOLUTIONAL NEURAL NETWORKS FOR LVCSR," in Acoustics, Speech and Signal Processing, 2013.

[14] B. Hutchinson, L. Deng and D. Yu, "A DEEP ARCHITECTURE WITH BILINEAR MODELING OF HIDDEN REPRESENTATIONS: APPLICATIONS TO PHONETIC RECOGNITION," in ICASSP 2012 - 2012 IEEE International Conference on Acoustics, Speech and Signal Processing, Kyoto International Conference Center (KICC), 2012.