Introduction to Deep Learning

Anjali D.Dwivedi
Student, M.E. Computer Engineering
Shree L.R.Tiwari College of engineering, Mumbai University
Mumbai, India
danjali319@gmail.com

K.jayamalini
Assistant Professor, Computer Engineering
Shree L.R.Tiwari College of engineering, Mumbai University
Mumbai, India
malini1301@gmail.com

Abstract— deep learning is a topic that has seen considerable media attention over the last few years. Deep learning is successful because it learns by a hierarchical system of features that bears similarity to the human mind. In this article we will introduce about deep learning, what is deep learning, why it is important its application and how it related to other techniques in the field of data science. We will also show how deep learning is different form machine learning and neural network.

Keywords—*Deep learning; neural network; machine learning;*

*****

I. INTRODUCTION

The goal of Artificial intelligence is to develop a machine which can remember, sense, learn, and recognize like a real human being. Perception is the first machine which can sense and learn but has fundamentally limited learning abilities. The later neural network [1] with multiple hidden layers can learn more complicated functions but it lacks a good learning algorithm. The appearance of Support Vector Machine enlighten people within a short time since it facilitates the learning procedures and performs well in many practical problems, but SVM also encounters its bottlenecks due to its shallow architectures. Feedback is part of the interactive components of teaching and learning and can therefore be seen as central to pedagogy. There are many ways in which teachers assist student by giving them feedback about their work for development of their learning skill. The important issue is that whatever the selected method, it should be able to provide information about what the student knows and does not know, as well as providing direction for improvement. Feedback can be provided on an individual and group basis.

Deep-learning is impacting everything from manufactu-ring, healthcare, transportation and many more. Companies are switching to deep learning to solve problems, like object recognition, speech recognition and machine translation. One of the most magnificent achievements this year was 'AlphaGo' a computer program that thrash the best Go player in the world. With the victory, Go joins checkers, othello, Jeopardy and chess are machine game in which machine have defeated human. While beating someone at a board game does not seem useful on the surface, this is a big deal. Before the victory, Go was written for a competent AI. Due in part to human intuition essential to play the game. The success makes an entire class of problems once considered intractable ready for solving. We are beginning to see commercial use. Such is the case with self-driving cars. Companies like Tesla, Google and Uber testing autonomous car on the streets.

As inferred "Deep learning is learning that takes root in our apparatus of understanding, in the embedded meanings that define us and that we use to define the world".

II. LITERATURE SURVEY

Survey describe that deep learning is becoming a mainstream technology for speech recognition at industrial scale. In this paper, we provide an overview of the work by Microsoft speech, since 2009 researcher in this area focusing on more recent advances which shed light to the basic capabilities and limitations of the current deep learning technology. We organize this view along with feature-domain and model-domain dimensions according to the conventional approach to analyzing speech systems. Selected experiments results, including speech recognition and related applications such as language modeling and spoken dialogue are presented to demonstrate and analyze the strengths and weakness of the techniques described in the paper.

It [2] presents that deep learning systems have dramatically improved the accuracy of speech recognition and various deep architectures and learning techniques have been developed with distinct strengths and weaknesses in recent years. How can ensemble learning be applied to the various deep learning systems to achieve greater recognition accuracy is the focus of this paper. We develop and report linear stacking methods for ensemble learning with applications specifically to speech-class and long-linear stacking methods for ensemble learning with applications connected deep neural networks.

Survey presents Visual speech information from the speakers moth region has been successfully shown to improve noise robustness of automatic speech recognizers, thus promising to increase their usage into the human computer interface. In this paper, we can see the main components of audio-visual automatic speech recognition and present novel contributions in two main areas: first, the visual front end design and later, we discuss new work on features and design fusion combination , the modeling of audio-visual speech asynchrony and incorporating modality reliability estimates to the bimodal recognition process.

It [3] presents a brief survey on speech is the primary and the most convenient means of communication between people. Speech has potential of being efficient mode of interaction with computer. This paper gives an overview of appreciation of the fundamental progress of speech recognition and major technological perspective

also gives overall idea about technique developed at every stage of speech recognition.

[4]This paper helps in selecting the technique along with their relative advantages and disadvantages. A comparative study of different techniques is done. This paper concludes with the decision on feature direction for development of technique in human computer interface system in distinct mother tongue and it also gives the various technique used in each step of a speech recognition process and try to analyze methods for designing an important system for speech recognition. The objective of this paper is to compare summarized different speech recognition systems and identify research topics and applications where are at the front end of this exciting and challenging field.

III. WHAT IS DEEP LEARNING

To understand what deep learning is, we first need to understand the relationship between deep learning and other machine learning technique like neural network, machine learning and artificial intelligence.

Lets us visualize figure1 to make understanding about their relationship

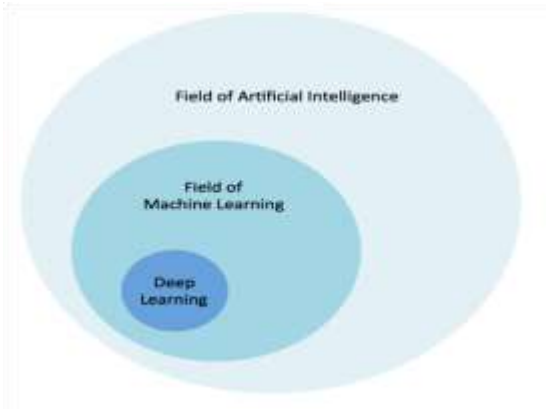

Fig. 1. Relationship between deep learning and other machine learning technique

At the outer most circles we have artificial intelligence. One circle inside of that is machine learning. With deep learning at centre. Deep learning means depth of the network i.e. artificial neural network, deep learning is easy and approachable name for artificial neural network is very shallow. Neural networks are based on the concept of cerebral cortex. As there are several layers of interconnected perceptrons in cerebral cortex likewise there are many layers of interconnected nodes.

The figure 2 shows the flow of deep learning, the only difference between neural network and deep learning is presence of 'hidden layer' in deep learning, first layer is always the input layer. Each node present in this layer takes an input, and then passes its output as the input to other nodes in the next layer. There are no connections between nodes in the same layer and the last layer always produces the outputs

We call the middle as the hidden layer. These neurons have no connection to the outside layer (e.g. input or output) and they are only activated by nodes in the previous layer.

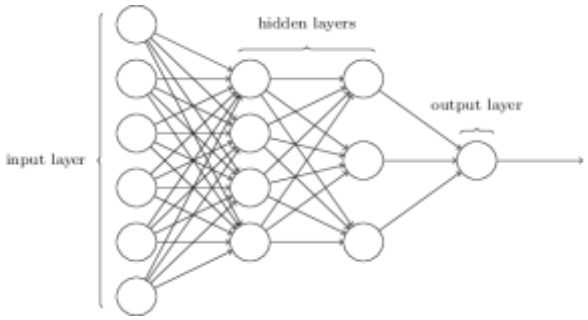

Fig. 2. Layers in deep learning

Deep learning is a technique for machine learning that utilizes multiple layer of abstraction for solving problem like pattern recognition. In the 1980s, most of the neural networks work on single layer due to the availability of data and cost of computation. Machine learning is considered a branch of Artificial intelligence, whereas deep learning is a type of machine learning. Machine learning involves computer intelligence that run against training data set and checks its success attempt and modify accordingly. Machine learning typical requires a spanning software engineering, sophisticated education and computer science to linear algebra and statistical methods.

Machine learning is divided into two broad categories:

- Unsupervised learning
- Supervised learning

In supervised learning, a machine learning algorithm uses a dataset with label to produce the desired outcome. This is time consuming as it takes a lot of data, and the data needs to be labeled manually. Supervised learning is best used for regression and classification problems. For example, let's say that we were running a firm and we want to know about effect of bonus on employee retention. If we had historical data i.e. employee bonus amount and tenure we could use supervised machine learning.

With unsupervised learning there aren't any corresponding or predefined answers. The goal is to find out the hidden patterns in the data. It's usually used for associative and clustering tasks, like grouping customers by what the purchase or behavior. For example Amazon's "customers who also bought ..." recommendations are a type of associative task. While supervised learning can be more useful, we often have to utilize unsupervised learning. Deep learning has proven to be an effective unsupervised learning technique.

IV. WHY IS DEEP LEARNING IMPORTANT

Computers have techniques for recognizing features inside of an image long ahead but the results weren't always appropriate. Deep learning was of great benefit for computer vision. [5] Computer vision by using deep learning now select humans on many image recognition tasks. Facebook had a great success with recognizing faces in photographs by using deep learning technique. It's not just a average improvement, but a game changer: "Asked whether two different photos of faces shows the same person, a human being will get it right 97.53 percent of the time. New software developed by researchers at Facebook using deep learning can score 97.25 percent on the same challenge, despite of variations in whether the person in the image is directly facing the camera or lightning."

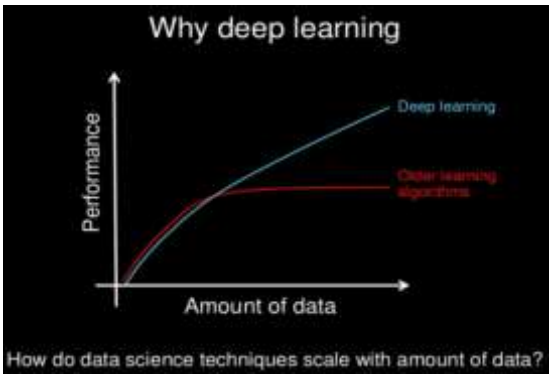

Fig. 3.data science technique scale with amount of data

Deep learning also impacted in other area i.e. Speech recognition. Languages spoken are so ambiguous and vast. Baidu – china's leading search engines – has developed a voice recognition system that is more accurate and faster than humans at generating text on a mobile phone. In both Mandarin and English. What is particularly interesting is that the two languages didn't require much design effort: "previously, people viewed English and Chinese as two different languages, and so there was a need to design very different features," chief scientist Andrew Ng says at Baidu. "The learning algorithms are now so general that you can just learn."Google is now using deep learning to manage company data centers. They've cut 40% of their energy needs for cooling, that translates to about improvement in power usage efficiency by 15% for the company and hundreds of millions of dollars in savings.

V. TRAINING DEEP NETWORK

5.1. Build a feature space

Note that this is what we do with trained hidden layers and Support vector machine kernels in BackPropogation, etc., but now we are going to build the feature space using deep architectures. [6] Unsupervised training between different layers can minimize the problem into distributed sub-problems to be further decomposable at subsequent layers for more precisely training the network.

5.2. Greedy Layer-Wise Training

Greedy layer-wise training ignores many of the problems of trying to train in a supervised fashion. Each layer gets full learning focus as it is the only current "top" layer can take advantage of unlabeled data. When you finally start the entire network with supervised training the network weights have already been adjusted so that you are in a good error basin and just need fine tuning.

- Steps of greedy layer-wise training

1. Train first layer without the label using your data (unsupervised) since there is no target at this level, labels don't help. Nodes could also use the more large unlabeled data which is not part of the training data set (i.e. self-taught learning).

2. Then freeze the first layer nodes or parameters and by using the output of first layers start training the second layer as the unsupervised input to the second layer.

3. Repeat this for as many layers in the network. This builds our set of strong features.

4. The outputs of the final layer are used as input to a supervised layer/model and train the last supervised layer(s) (freeze the early weight).

5. Release/unfreeze all weights and tune the full network by training with a supervised approach, given the pre-training weight settings.

VI. DEEP NEURAL ARCHITECTURE

6.1. Convolution Neural Network

A CNN is composed of one or many convolutional layers with fully connected layers on top. It also uses pooling layers and tied weight. The concept of max-pooling is used by it. CNN have shown efficient results in both speech and image applications. They can also be trained with back propagation. CNNs are easier to train making them a highly attractive architecture to use. Examples of applications in Computer Vision include Deep Dream.

6.2. Recursive neural network

A recursive neural network is developed by applying the same set of weights recursively over a different graph-like structure. Such networks are trained by the reverse mode of autonomous differentiation. They were introduced to learn such as logical terms. A better case of recursive neural networks is the RNN itself whose structure corresponds to a linear chain. Recursive neural networks have been applied to natural language processing.

6.3 Convolutional deep belief networks

[7] A recent achievement in deep learning is the use of Convolutional deep belief networks (CDBN). CDBNs have structure very similar to a CNN and are trained similar to DBN (deep belief network). Therefore, they exploit the 2D structure of images, like CNNs do, and make use of pre-training like deep belief networks. They provide a generic structure which can be used in many image and signal processing tasks. Recently, many effective results on CIFAR a standard data set have been obtained using CDBNs.

VII. OPEN SOURCE DEEP LEARNING FRAMEWORKS

Deep learning's tool is made accessible by a no. of open source projects. Some of the most used technologies include Deeplearning4j (DL4j), Theano, Torch, TensorFlow, and Caffe. The deciding factors on which framework to use is the tech stack they target, and if they are application focused or low-level academia. Here's an overview of each:

DL4j:

- JVM-based.
- Distrubted.
- Integrates with Spark and hadoop.

Theano:

- Most popular in Academia.
- Fairly low level.
- Interfaced with via Python and Numpy.

Torch:

- Lua based.
- In house versions used by Facebook and Twitter.
- Contains pertained models.

TensorFlow:

- Google written successor to Theano

- Interfaced with via Python and Numpy.
- Highly parallel.
- Can be bit slowly for certain problem sets.

Caffe:

- Not general purpose.
- Focuses on machine-vision problems.
- Not easily extensible.
- Implemented in C++ and is very fast.
- Has a Python interface.

VIII. APPLICATION

Deep learning has many practical applications that company and businesses are using today,[11] and many more that will be used as research continues. Popular uses today include:

1. Speech Recognition

Both the academic and business worlds have accepted deep learning for speech recognition. Skype, Xbox, Apple's Siri and Google Now, to name a few, are already using deep learning technologies in their systems to identify voice pattern and human speech.

2. Image Recognition

One practical application of image recognition is scene description and automatic image captioning. This could be beneficial in law enforcement investigations for recognizing criminal activity in thousands of photos submitted by person in a crowded area where a crime has taken place. Self-driving cars will also take advantage of image recognition by using 360-degree camera technology.

3. Natural Language Processing

Analyze written text for many years. A specialization of text mining, this technique can be used to discover patterns in customer complaints, news reports or physician notes, to name a few.

4 Recommendation Systems

Netflix and Amazon have popularized the usage of a recommendation system with a good knowledge about what you might be interested in next, based on past behavior. Deep learning can be used to enhance recommendations in complex environments such as clothing preference and music interests across multiple platforms.

IX. COMPARATIVE STUDY BETWEEN DEEP LEARNING AND OTHER MACHINE LEARNING TECHNIQUE

Deep learning is a basic practice to parse data using algorithm ,learn from it and then make a prediction about things in the world, so rather than coding we train our machine to accomplish task using large amount of dataset to trained itself.

Deep Learning is broader family of Machine Learning [8] methods that tries to learn high level features from the given data. Thus, the problem it solves is reducing task of making new feature extractor for each and every type of data (speech, image etc.) for example deep learning algorithm will try to learn about features such as length of nose, distance between eyes etc. in future when image recognizition task is given to them they may use this information for prediction and classification etc this is major step away from previous shallow learning algorithm for example if we write below equation F(1,2,3.....,100)=5050
and give it to a ML Algorithm, then that Algorithm is like the child which immediately understands that RHS is sum of all numbers in LHS. Given new set of numbers F(1,2,3.....,500), the child will then sum up all the numbers up to 500.Then, Deep Learning Algorithm will realize that sum of **i**th element from forward and backward is always fixed, and can be used to find the sum.

Machine learning is the field of making computers smarter and able to learn from the data rather than static instructions. Whereas Deep learning is a sub-field of machine learning focused on high-level abstracted representation of data using multiple processing layers (thus the deep qualifier).

Deep Learning is about constructing machine learning models that learn a hierarchical representation of the data.

Neural Networks are the artificial neuron forms the computational unit of the model and the network describes how these units are connected to [9] one another. You can describe a hierarchical model with Neural Networks where each layer of neurons represents a level in that hierarchy. There is no real restriction in how many layers you can add to a network but going beyond two layers, was impractical in the past with diminishing returns. The limitations were overcome with algorithmic advances, availability of large volumes of training data and accelerated GPU computing. People started adding more layers again, resulting in Deep Neural Networks with much success. They demonstrated how Neural Networks with lots of layers enabled constructing the representations required for deep learning.

Deep Neural Networks are the showcase for Deep Learning. Some argue that Decision trees and Hierarchical temporal memory qualify as non-neural deep learning models. The point is that the concept of Deep Learning is not tied to Neural Networks but is very well demonstrated through them.

X. SCOPE AND CHALLENGES

The new challenges of deep learning was making intelligent machines capable of not only seeing (vision) and hearing (speech), but also of thinking from "brain"; i.e. inference and reasoning over hierarchical complex knowledge and relationships sources that consist a vast number of semantics and entities concepts in the real world based on multi-sensory data from the user. [10] To this end, multimodal and language processing, learning from speech/audio, text and image/video, joint exploitation is evolving into a new features of deep learning, beginning to be accepted by a research communities including spoken language processing and speech recognizing, natural language processing, machine learning, computer vision, cognitive science information retrieval, artificial intelligence, and knowledge management.

A review of recently published studies related to deep learning will be applied to multimodal processing task and selected language with a trace back to neural network literature and with future directions in this new exciting deep learning frontier analyzed and discussed.

XI. CONCLUSION AND FURTURE SCOPE

We have introduced a basis introduction about what is deep learning, how it is different from other machine learning technique. We know that DL provide great learning feature as compared to other machine learning. Algorithm and architecture are main component of DL (Deep Learning). Architectures like DBNs and CNNs perform well in many AI tasks.

Is it possible to implement deep learning without the deep architectures? A recent work Saul and Cho who come from UCSD prove that kernel machines can be used for deep learning. The method they use is to apply multiple times of feature mapping to enact the computation of deep learning. They apply feature mapping method to solve the image recognition problem, which performs better than the DBNs and SVM with Gaussian kernel.

This work gives us a new way in exploring deep learning, which also provide the fact that the deep architecture is proved to be a good model for the deep learning, but not the best one. There might be many things to explore in this amazing field.

REFRENCES

[1] Dandan Mo, L. Stochastic gradient learning in neural networks: A survey on deep learning: one small step toward AI 2012.

[2] Xavier Glorot, Antoine Bordes et. al : Domain Adaptation for Large-Scale Sentiment Classication:A Deep Learning Approach, 2008.

[3] ALISON RUSHTON, A. University of Birmingham, Edgbaston,Birmingham, UK : Formative assessment: a key to deep learning? , 2007.

[4] Ciresan, D. C., Meier, U., Gambardella, L. M., and Schmidhuber, J. Deep big simple neural nets excel on handwritten digit recognition. CoRR, 2010.

[5] Coates, A., Lee, H., and Ng,A. Y. An analysis of singlelayer networks in unsupervised feature learning. In AISTATS 14, 2011.

[6] Dean, J. and Ghemawat, S. Map-Reduce: simplified data processing on large clusters. Comm. ACM, pp. 107–113,2008.

[7] Deng , Li,Gong,et al. : Proximal regularization for online and batch learning. In ICML, 2009.

[8] Good fellow, I., Le, Q.V., Saxe, A., Lee, H., and Ng, A.Y. Measuring invariance's in deep networks. In NIPS, 2010.

[9] Gaikwad, Gawali and Yannawar : A practical guide to training restricted boltzmann machines. Technical report, U. of Toronto, 2010.

[10] Lee, H., Ekanadham, C., and Ng, A. Y. Sparse deep beliefnet.

[11] Nidhi Bhatia, Mrs.ChhaviRana :Deep learning technique and its various application. In IJEIR 2015