International Journal of Engineering Innovation & Research
Volume 4, Issue 5, ISSN: 2277 – 5668

# Deep Learning Techniques and its Various Algorithms and Techniques

**Nidhi Bhatia, Mrs. Chhavi Rana**
University Institute of Engineering and Technology, M.D.U. , Rohtak

*Abstract* – **The concept of deep learning is not new to higher education. However, deep learning has drawn more attention in recent years as institutions attempt to tap their student's full learning potential. To more fully develop student talents, many campuses are shifting from a traditional passive, instructor- dominated pedagogy to active, learner-centered activities. Switching these features of human brain to a learning model, we wish the model can deal with the high-dimensional data, support a fast and intellectual learning algorithm and perform well in the complicated AI tasks like computer vision or speech recognition. This survey reviews a history of deep learning, summarizing the components of Convolutional Neural Networks(CNNs) and Deep Belief Networks (DBNs) together with their learning algorithms and their performances in different applications. Institutions and researchers can use the resulting scales to assess and investigate deep approaches to learning.**

*Keywords* – **Deep Learning, Deep Belief Networks (DBNs), Convolutional Neural Networks (CNNs), Algorithm.**

## I. Introduction

The goal of AI is to invent a machine which can sense, remember, learn, and recognize like a real human being. Perceptron 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 SVM enlightens 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[2]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 can provide feedback to assist the development of student learning. The important issue is that whatever the selected method, it must be able to provide information about what the student does and does not know, as well as providing direction for improvement. Feedback can be provided on an individual and group basis.

Students using "surface-level processing[3]"focus on the substance of information and Emphasize rote learning and memorization techniques. The goal of studying for a test or exam is to avoid failure, instead of grasping key concepts and understanding their relation to other information and how the information applies in other circumstances. In contrast, students using "deep-level processing" focus not only on substance but also the underlying meaning of the information. As inferred: "Deep learning[4]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".

### *1.1 Learning*

Learning[5] can be considered to encompass 'deep learning' that includes understanding and interpretation. It is recognized the potential that assessment has for affecting learning and the intricate links are now widely recognized informing pedagogy. In the literature formative assessment is linked more strongly to teaching rather than learning. The literature highlights many claims regarding the positive effects of formative assessment on learning, Further work applying the existing theories into practice is therefore necessitated. There are many aspects of classroom interaction that contribute to formative assessment, such as discourse, questioning, giving tests and observation. In studies, it was found that feedback was the greatest influence on performance if provided prior to provision of the answers. There is considerable literature addressing this area, but there is considerable variation between the existing studies that limits the internal validity of using such meta-analyses to inform practice.

## II. Literature Review

Deng, Li,Gong [1]: 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 researchers since 2009 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 spoken dialogue and language modeling are presented to demonstrate and analyze the strengths and weakness of the techniques described in the paper.

Deng, Platt[2]: survey 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 ensembling learning with applications specifically to speech-class and long-linear stacking methods for ensemble learning with applications connected deep neural networks.

International Journal of Engineering Innovation & Research
Volume 4, Issue 5, ISSN: 2277 – 5668

Gravier, Garg[3,10] : survey presents Visual speech information from the speaker's moth region has been successfully shown to improve noise robustness of automatic speech recognizers , thus promising to extend their usability into the uman computer interface. In this paper, we review 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.

Das[4]: presents a brief survey on speech is the primary and the most convenient means of communication between people. The communication among human computer interaction is called human computer interface. Speech has potential of being important mode of interaction with computer. This paper gives an overview of major technological perspective and appreciation of the fundamental progress of speech recognition and also gives overview technique developed in each stage of speech recognition. This paper helps in choosing the technique along with their relative merits and demerits. A comparative study of different techniques is done. This paper concludes with the decision on feature direction for developing technique in human computer interface system in different mother tongue and it also gives the various technique used in each step of a speech recognition process and attempts to analyze an approach for designing an efficient system for speech recognition . The objective of this review paper is to summarize and compare different speech recognition systems and identify research topics and applications where are at the front end of this exciting and challenging field.

Dhameliya, Desai[7,8]: survey presents speech is the most natural form of human communication and speech processing has been one of the most inspiring expanses of signal processing . Speech recognition is the process of automatically recognizing the spoken words of person based on information in speech signal. Automatic Speech Recognition(ASR) system takes a human speech utterances as an input and requires a string of words as output. This paper introduces a brief survey on Automatic Speech Recognition and discusses the major subjects and improvements made in the past 60 years of research , that provides technological outlook and a respect of fundamental achievements that have been accomplished in the important areas of speech recognition. Definition of various types of speech classes , feature extraction techniques, speech classifiers and performance evaluation are issues that require attention in designing of speech recognition system. The objective of this review paper is to summarize some of the well-known methods used in several stages of speech recognition system.

Gaikwad, Gawali and Yannawar [6,9] : The speech is most prominent and primary mode of communication among human beings. The communication among human computer interaction is called human computer interface. Speech has potential of being important mode of interaction with computer. This paper gives an overview of major technological perspective and appreciation of fundamental progress of speech recognition and also gives an overview technique developed in each stage of speech recognition. This paper helps in choosing the technique along with their merits and demerits. A comparative study of different techniques is done. This paper concludes with the decision on feature direction for developing technique in human computer interface system in Marathi language.

Therese, Lingam [5,11]: Says that speech has evolved as a primary form of communication between humans. The advent of digital technology gave us highly versatile digital processors with high speed, low cost and high power, which enable researchers to transform the analog speech signals into digital speech signals that can be significantly studied. Achieving higher recognition accuracy, low word error rate and addressing the issue of resources of variability are the major consideration for developing an effective automatic Speech Recognition System. . In speech recognition, feature extraction requires much attention because recognition performance depends heavily on this phase. In this paper, an effort has been made to highlight the progress made so far in the feature extraction phase of spec recognition system and an overview of technological perspective of Automatic Speech Recognition System is discussed.

## III. TECHNIQUES

*Deep learning algorithms*

*(a) Restricted Boltzmann Machines*

In RBMs, the gradient used in training is an approximation formed by a taking small number of Gibbs sampling steps. Given the biased nature of the gradient and intractability of the objective function, it is difficult to use any optimization methods other than plain SGDs.

*(b) Auto encoders and denoising auto encoders*

Here, we use the L2 norm to penalize the difference between the reconstruction and the input. Typically, we set the activation function _ to be the sigmoid or hyperbolic tangent function. Unlike RBMs, the gradient of the auto encoder objective can be computed exactly and this gives rise to an opportunity to use more advanced optimization methods, such as L-BFGS and CG, to train the networks.

*(c) Sparse RBMs and Auto encoders*

Sparsity regularization typically leads to more interpretable features that perform well for classification. Sparse coding was first proposed by (Olshausen & Field, 1996) as a model of simple cells in the visual cortex. The key idea in this approach is to penalize the deviation between the expected value of the hidden representations and preferred target activation _. By setting _ to be close to zero, the hidden unit will be sparsely activated. Sparse representations have been employed successfully in many applications such as object recognition, speech recognition and activity recognition.

A common practice to train sparse RBMs is to use a running estimate and penalizing only the bias. This further complicates the optimization procedure and makes it hard to debug the learning algorithm. Moreover, it is important to tune the learning rates correctly for the different parameters W, b and c. Consequently, it can be difficult to train sparse RBMs.

In our experience, it is often faster and simpler to obtain sparse representations via auto encoders with the proposed sparsity penalties, especially when batch or large mini batch optimization methods are used.

In detail, we consider sparse auto encoders with a target activation of _ and penalize it using the KL divergence.

To train sparse auto encoders, we need to estimate the expected activation value for each hidden unit. However, we will not be able to compute this statistic unless we run the optimization method in batch mode. In practice, if we have a small dataset, it is better to use a batch method to train a sparse auto encoder because we do not have to tweak optimization parameters, such as mini batch size, _ as described below.

*(d) Tiled and locally connected networks*

RBMs and auto encoders have densely-connected network architectures which do not scale well to large images. For large images, the most common approach is to use convolutional neural networks. Convolutional neural networks have local receptive field architectures: each hidden unit can only connect to a small region of the image. Translational invariance is usually hardwired by weight tying. Recent approaches try to relax this constraint.

It shows that local architectures, such as tiled convolutional or convolutional architectures, can be efficiently trained with a computer cluster using the Map-Reduce framework. With local architectures, the cost of communicating the gradient over the network is often smaller than the cost of computing it (e.g., cases considered in the experiments).

## IV. Challenges

1. The previous protocol exhibits appealing properties for domain adaptation of sentiment classifiers. Existing domain adaptation methods for sentiment analysis focus on the information from the source and target distributions, whereas the SDA unsupervised learning can use data from other domains, sharing the representation across all those domains.
2. This also reduces the computation required to transfer to several domains because a single round of unsupervised training is required, and allows us to scale well with large amount of data and consider real-world applications.
3. The code learned by the SDA is a non-linear mapping of the input and can therefore encode complex data variations.
4. To the best of our knowledge, existing domain adaptation methods for sentiment analysis map inputs into a new or an augmented space using only linear projections.
5. Furthermore, rectifier non-linearity’s have the nice ability to naturally provide sparse representations (with exact zeros) for the code layer, which are well suited to linear classifiers and are efficient with respect to computational cost and memory use.

## V. Conclusion

The results of this study suggest that these items, when combined with existing core survey items, assess three distinct aspects of a second order factor that, in content, appears to be related to deep learning. Natural Language Processing(NLP) is a typical example; deep learning cannot understand a story, as well as a general request to an expert system. So there's still a long way to go before we can implement the real intelligent machine. But deep learning indeed provides a direction to implement the more intellectual learning; therefore it can be regarded as a small step toward AI.

- Deep architectures help deep learning by trading a more complicated space for better performance, in some cases, even for less computation time.
- Deep architectures are good models for deep learning, but can't be proved to be the best one. There're still many possibilities in the architectures and learning algorithms that can carry out better performances.

## Future Scope

From the analysis above, we know that deep learning represents a more intellectual behavior(learning features) compared with the other traditional machine learning. Architectures and the related learning algorithms are the two main components of deep learning. From the analysis above, we know that deep architectures like CNNs and DBNs perform well in many AI tasks. But is it true that only deep architectures can implement deep learning? Is it possible to implement deep learning without the deep architectures? A recent work Cho and Saul who come from UCSD shows that kernel machines can also be used for deep learning. The approach they use is to apply multiple times of feature mapping to mimic the computation of deep learning. They apply this method to solve the image recognition problem, which performs better than the SVM with Gaussian kernel as well as the DBNs. This work gives us a new direction in exploring deep learning, which also indicates 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 surprises waiting for us to explore in this amazing field.

## Comparative Study

In the original paper regarding the smaller 4-domain benchmark dataset adapt Structural Correspondence Learning(SCL) for sentiment analysis. Li and Gong [7] (2009) propose the Multi-label Consensus Training(MCT) approach which combines several base classifiers trained with SCL. Pan et al. (2010) first use a Spectral Feature Alignment (SFA) algorithm to align words from different source and target domains to help bridge the gap between them. These 3 methods serve as comparisons in our empirical evaluation.

## REFERENCES

[1] Das, Thomas F. Nelson Laird et al. : Measuring Deep Approaches to Learning Using the National Survey of Student Engagement, 1900
[2] Dandan Mo, L. Stochastic gradient learning in neural networks: A survey on deep learning: one small step toward AI 2012.
[3] Xavier Glorot, Antoine Bordes et. al : Domain Adaptation for Large-Scale Sentiment Classi_cation:A Deep Learning Approach, 2008.
[4] ALISON RUSHTON, A. University of Birmingham, Edgbaston,Birmingham, UK : Formative assessment: a key to deep learning? , 2007.
[5] Ciresan, D. C., Meier, U., Gambardella, L. M., and Schmidhuber, J. Deep big simple neural nets excel on handwritten digit recognition. CoRR, 2010.
[6] Coates, A., Lee, H., and Ng,A. Y. An analysis of singlelayer networks in unsupervised feature learning. In AISTATS 14, 2011.
[7] Dean, J. and Ghemawat, S. Map-Reduce: simplified data processing on large clusters. Comm. ACM, pp. 107–113,2008.
[8] Deng , Li,Gong,et al. : Proximal regularization for online and batch learning. In ICML, 2009.
[9] Good fellow, I., Le, Q.V., Saxe, A., Lee, H., and Ng, A.Y. Measuring invariances in deep networks. In NIPS, 2010.
[10] Gaikwad, Gawali and Yannawar : A practical guide to training restricted boltzmann machines. Technical report, U. of Toronto, 2010.
[11] Lee, H., Ekanadham, C., and Ng, A. Y. Sparse deep beliefnet model for visual area V2. In NIPS, 2008.

## AUTHOR'S PROFILE

**Nidhi Bhatia**
B.Tech., M.Tech., M. D. University, Rohtak