LITERATURE REVIEW OF DEEP LEARNING TECHNIQUES

Payal Singh, Amit Pratap Singh, Dr. Anshuman Tyagi

Student, Assistant Professor, Head Of Department

Electrical & Electronics Engineering

Pranveer Singh Institute of Technology, Kanpur, Uttar Pradesh, India

Abstract- Deep learning is a part of machine learning, which aims to learn a hierarchy of features from input data. Researchers have nowadays intensively investigated deep learning algorithms for solving challenging problems in many areas such as image classification, speech recognition, signal processing, and natural language processing.

Keywords:- Deep Learning; Techniques

I.   INTRODUCTION

The goal of AI is to invent a machine that can sense, remember, learn, and recognize like human being. Perceptron is known to be the first machine which can sense and learn but has fundamentally limited learning abilities. The later neural networks containing multiple hidden layers can learn more complicated functions but lacks a good learning algorithm. The appearance of SVM enlightens people within a short time span since it facilitates the learning procedures and performs well in many practical problems, but SVM also encounters its limits due to its shallow architectures. Feedback is part of the interactive components of teaching and learning and can thus be seen as central to pedagogy. There are various ways in which teachers can provide feedback to assist the development of student's learning. The important issue is that whatever the selected method, it must be able to provide an accurate 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 on group basis. Students using "surface-level processing" focus on the substance of information and Emphasize on some rote learning and memorization techniques. The goal of studying for a test or exam is just to avoid failure, instead of grasping key concepts and understanding their relation to other information and how the information applies in various other circumstances. In contrast, students using "deep-level processing" focuses not only on substance but also the meaning of the information. As inferred: "Deep learning is learning that takes root in our apparatus of understanding, in the embedded meanings that define us and the way we use to define the world".

1.1 Learning

Learning can be considered to include deep learning" which includes understanding and interpretation." It is recognized the potential that assessment has for affecting learning and the intricate links are now widely recognized as informing pedagogy. In the literature formative assessment is linked more strongly to teaching rather than learning. The literature highlights may claims regarding the positive effects of formative assessment on learning, Further works applying the existing theories into practice is therefore necessitated. There are many aspects of classroom interaction which contributes to formative assessment, such as discourse, questioning, giving tests and observation. Studies discovered 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 difference between the existing studies that limits the internal validity.

II. LITERATURE REVIEW

Deng, Li,Gong : The survey describe that deep learning is becoming a mainstream technology for speech recognition at industrial scale. In this paper, we overview the work by Microsoft speech researchers since 2009 in this area and focus on various developments and shed light on future advancements and limitations of the current deep learning technology. We organize this along with feature-domain and model-domain dimensions of the conventional approach to analyze speech system.

Deng, Platt: survey presents that deep learning systems have improved the accuracy of speech recognition and various deep architectures and learning techniques have been developed in recent years.

Gravier, Garg : survey presents Visual speech information from the speaker's mouth region has been successfully shown to improve noise robustness of automatic speech recognizers. In this paper, the main components of audio-visual automatic speech recognition is reviewed and novel contributions in two main areas is presented: first, the visual front end design and later, we discuss new work on features and design fusion combination.

Das: presents a survey on speech is the primary and the most convenient means of communication between people. The communication among human computer interaction is known as human computer interface. The 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. The paper helps in choosing the techniques along with their relative merits and demerits. A comparative study of different techniques is done. The paper concludes with the decision on feature direction for developing techniques 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 .

Dhameliya, Desai: survey presents speech as the most natural form of human communication and speech processing being 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 as output. The paper introduces a brief survey on Automatic Speech Recognition and discusses the major changes made in the past 60 years of research , that provides technological outlook and a respect of fundamental achievements that have been accomplished in important areas of speech recognition. Definition of various types of speech classes , feature extraction techniques, speech classifiers and performance evaluation are some issues that require attention in designing of speech recognition system.

Therese, Lingam : Says that speech has evolved as a primary form of communication among humans. The advent of digital technology gave us highly versatile digital processors with high speed, low cost and high power, that enable researchers to transform the analog speech signals into digital speech signals that can be significantly studied. Achieving higher recognition accuracy, lower 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 takes much attention because recognition performance depends heavily on this phase. In the paper, an effort has been made to highlight the progress made so far in the feature extraction phase of speech recognition system and an overview of technological perspective of Automatic Speech Recognition System is discussed.

III. TECHNIQUES
3.1. Deep learning algorithms

(a) Restricted Boltzmann Machines
It was invented by Geoffrey Hinton.A Restricted Boltzmann machine is an algorithm useful for dimensionality reduction, classification, regression, collaborative filtering, feature learning and topic modeling. In this, the gradient used in training is an approximation formed by a taking small number of Gibbs sampling steps.If the biased nature of the gradient and intractability of the objective function is given, it is difficult to use any optimization methods other than plain SGDs.

(b) Auto encoders and denoising auto encoders

Autoencoders are Neural Networks which are commonly used for feature extraction. However, when there are more nodes in the hidden layer than there are inputs, the Network is risking to learn the "Identity Function", also known "Null Function", meaning that the output same as the input, marking the Autoencoder useless. In this, we use the L2 norm to penalize the difference between the reconstruction and the input. We set the activation function _ to be the sigmoid or hyperbolic tangent function. Unlike in 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 leads to more interpretable features that perform well for classification. Sparse coding was initially 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. Further complicates the optimization procedure and makes it hard to debug the learning algorithm. Moreover, it becomes 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. 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 size images, the most common approach is to use convolutional neural networks. Convolutional neural networks have a 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 a 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. The 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 various 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.

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

VI. ACKNOWLEDGMENT

VII. REFERENCES

- Das, Thomas F. Nelson Laird et al. : Measuring Deep Approaches to Learning Using the National Survey of Student Engagement, 1900

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

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

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

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

- Coates, A., Lee, H., and Ng,A. Y. An analysis of single layer networks in unsupervised feature learning. In AISTATS 14, 2011.

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

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

- Good fellow, I., Le, Q.V., Saxe, A., Lee, H., and Ng, A.Y. Measuring invariances in deep networks. In NIPS, 2010.

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

- Lee, H., Ekanadham, C., and Ng, A. Y. Sparse deep beliefnet model for visual area V2. In NIPS, 2008.