International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 17
November 2016

Deep Belief Networks Using Convolution Neural Networks Algorithm

OGGU UMARANI[1] & MS.DAMALLA JYOTHI[2]

[1]M-Tech Dept of CSE, Laqshya Institute of Technology and Science, Khammam.

Mail id: umarani.banne@gmail.com

[2]Associate Professor Dept of CSE, Laqshya Institute of Technology and Science, Khammam.

Mail id: damallajyothi@gmail.com

Abstract

The concept of deep learning is not new to higher educatio n. 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), Convolution Neural Networks (CNNs), Algorithm.

1. 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.

2. Related Work

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.

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.

3. Implementation

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 hardto 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 convolution or convolution 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).

4. Experimental Work

We present the results from a series of experiments designed to test the effectiveness of our HF approach on the deep auto-encoder

International Journal of Research
IJR
International Journal of Research
Available at https://edupediapublications.org/journals
p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 17
November 2016

problems considered by Hinton & Salakhutdinov (2006) (abbr. H&S). We adopt precisely the same model architectures, datasets, loss functions and training/test partitions that they did, so as to ensure that our results can be directly compared with theirs. Each dataset consists of a collection of small grey-scale images of various objects such as hand-written digits and faces. Table 1 summarizes the datasets and associated experimental parameters, where size gives the size of the training set, K gives the size of minibatches used, and encoder dims gives the encoder network architecture. In each case, the decoder architecture is the mirror image of the encoder, yielding a “symmetric auto encoder”. This symmetry is required in order to be able to apply H&S’s pre-training approach. Note that CURVES is the synthetic curves dataset from H&S’s paper and FACES is the augmented Olivetti face dataset.

We implemented our approach using the GPU-computing MATLAB package Jacket. We also re-implemented, using Jacket, the precise approach considered by H&S, using their provided code as a basis, and then re-ran their experiments using many more training epochs than they did, and for far longer than we ran our HF approach on the same models. With these extended runs we were able to obtain slightly better results than they reported for both the CURVES and MNIST experiments. Unfortunately, we were not able to reproduce their results for the FACES dataset, as each net we pre-trained had very high generalization error, even before fine-tuning. We ran each method until it either seemed to converge, or started to over fit (which happened for MNIST and FACES, but not CURVES). We found that since our method was much better at fitting the training data, it was thus more prone to

Table 2. Results (training and test errors)

	PT + NCG	RAND+HF	PT + HF
CURVES	0.74, 0.82	0.11, 0.20	0.10, 0.21
MNIST	2.31, 2.72	1.64, 2.78	1.63, 2.46
MNIST*	2.07, 2.61	1.75, 2.55	1.60, 2.28
FACES	-, 124	55.4, 139	-,-
FACES*	-,-	60.6, 122	-,-

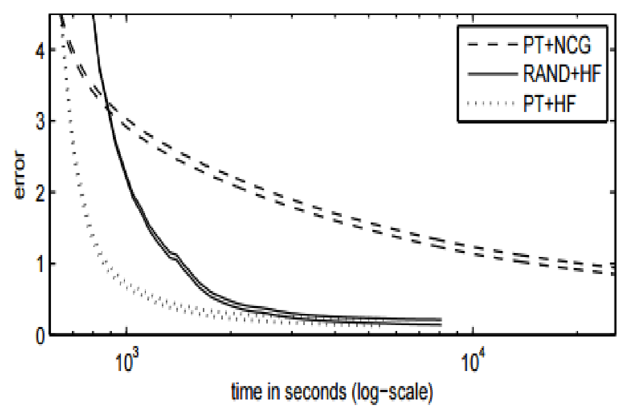

Figure 2. Error (train and test) vs. computation time on CURVES.

Over fitting, and so we ran additional experiments where we introduced an l2 prior on the connection weights. Table 2 summarizes our results, where PT+NCG is the pre-training + non-linear CG fine-tuning approach of H&S, RAND+HF is our Hessian-free method initialized randomly, and PT+HF is our approach initialized with pretrained parameters. The numbers given in each entry of the table are the average sum of squared reconstruction errors on the training-set and the test-set. The's indicate that an l2 prior was used, with strength 10−4 on MNIST and 10−2 on FACES. Error numbers for FACES which involve pre-training are missing due to our failure to reproduce the results of H&S on that dataset (instead we just give the test error number they reported). Figure 2 demonstrates the performance of our implementations on the CURVES dataset. Pre-training time is included where applicable. This plot is not meant to be a definitive performance analysis, but merely a demonstration that our method is indeed quite efficient.

5. 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.

6. References

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

[2] Dandong 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.

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 03 Issue 17
November 2016

[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.

OGGU UMARANI

B-Tech in Laqshya Institute Of Technology And Science, Khammam, M-Tech Computer Science And Engineering Laqshya Institute Of Technology and Science, Tanikella, Khammam.

Mail id: umarani.banne@gmail.com

Phone Number: 8008183643

Authors Profile

MS.DAMALLA JYOTHI

Designation: Associate Professor.

Working: Laqshya Institute Of Technology and Science,Tanikella, Khammam. Experience: 10years. Qualification: M.Tech.,(Ph.D)

Email-d: damallajyothi@gmail.com

Phone Number: 9885329435.