Архитектура нейронных сетей с глубоким обучением анализу прикладных задач

Орман Индира Маликовна

Евразийский национальный университет им. Л.Н. Гумилева, физико-технический факультет,
e-mail: indira.malikovna@mail.r

И.М. Орман1, С.А. Боранбаев1 И.Г.Курмашов1
1Евразийский национальный университет имени Л.Н.Гумилева,
Астана, Казахстан

Architecture of neural networks with deep learning to analyze an applied problem

Abstract

Deep Learning is a field included in to Artificial Intelligence. It allows computational models to learn multiple levels of abstraction with multiple processing layers. This Artificial Neural Networks gives state-of-art performance in various fields like Computer Vision, Speech recognition and different domain like bioinformatics. There are mainly three architectures of Deep Learning Convolution Neural Network, Deep Neural Network and Recurrent Neural Network which provides the higher level of representation of data at each next layer. Deep Learning is required to classify high dimensional data like images, audio, video and biological data.

Keywords: Neural Network, Deep Learning, Deep Neural Network, Stacked Autoencoder, Convolu- tion Neural Network, Recurrent Neural Network.

1. Introduction

Machine-Learning becomes vitally important in this era used for the decision making systems, Recommendation systems, to identify objects from the images, in web searching extra to make human beings life easier. Machine learning algorithms use training examples to uncover underlying patterns, build a model, and then build predictions on the new information based on the model. There is several of Machine Learning Algorithms used for Classification one of them is an Artificial Neural Network (ANN). Expertise shows that neural networks are superb pattern recognizers that even have the power to learn and build distinctive structures for a selected problem [13]. Ability to process natural data in their raw form is difficult task for Conventional machine-learning techniques [8]. If we construct machine learning system then we needed domain expertise and careful engineering to develop feature extractor that converts the raw information and create feature vector which feed to the input layer of the Neural Network. Perceptron and single hidden layer Neural Network requires handmade features as input where as Deep Neural Networks extract features by themselves with multiple hidden layers. Deep Learning also called as representation learning [3].

For example, in image recognition, it can be interpreted that feature learning is done in the order of pixel, edge, texton, motif, part, and object. Similarly, in text recognition, features are learned in the order of character, word, word group, Clause, sentence, and story[7]. Deep learning is additionally known as at representation learning, i.e., learning representations of the information that create it easier to extract helpful information once building classifiers or alternative predictors [3].

2. Deep Architectures

There are three categories of Deep Learning architectures namely Deep Neural Network, Convolution Neural Network and Recurrent Neural Network [10].

2.1 Deep Neural Network

DNN can be classified as MLP (Multilayer Perceptron), SAE(Stacked Autoencoder), and DBN(Deep Belief Network) based on the types of layer used in DNN and the corresponding learning method [10].

Multilayer Perceptron (MLP): MLP has more non-linear layers are stacked and is trained in a purely supervised manner that uses only data with labels by initializing the parameters randomly and then training with backpropagation algorithm and gradient. MLP is trained by large number of labelled data [10].

Deep Belief Network (DBN): DBN is generative models that are different than discriminative nature of conventional Neural Networks. It

Figure 1

DBN with multiple layers of RBM

provides joint probability distribution means P(label|Observations) and P(Observations|Label) over observable data and labels [2].

DBN is comprising two segments called as associative memory and stacked Restricted Boltzmann Machines(RBM). To learn DBN there are two phases called pre-training phase and fine-tuning phase. Pre-training phase is unsupervised learning with stacked RBMs. Backpropagation is needed to train the network. Wake-Sleep algorithm is used for weight updating if DBN is used as weight generative model [11].

Stacked Autoencoder (SAE): Autoencoder is the network which takes the original inputs and reconstructs the input which means the network is used for extracting the features which are related to application. So the

Figure 2

Autoencoder with inputs and reconstructed input as output

main goal of this network is try to regenerate the inputs by setting target values to be equal to inputs (i.e., it uses x^(i) = x(i), where x is an inputs and x^is output value) [7]. Architecture of Autoencoder shows in below figure

2. SAE is used for pre-training of deep network by "stacked" Autoencoder in greedy layer wise structure which makes network deeper. It train the network step-by-step that is output of first hidden layer becomes input for the next layer.

2.2 Convolution Neural Network

Convolution Neural Networks (CNN) is naturally motivated variations of MLPs. From Hubel and Wiesel's initial work on the visual cortex of feline, they found that this cortex contains complex course of arrangement of cells, this cells performs diverse tasks [10]. There are two sorts of cells called simple and complex cells which are associated with little part of the visual field. These little sub-areas of the visual field, called a receptive field. The sub-regions are move over the visual field to cover the whole visual field.

CNN are used for extracting patterns form the multi array type of data [8] [10] (i.e. high dimensional data- too many features available for the classification task) like 2D for images, 1D signals and 3D for video. The basic concepts behind CNN are Local Connectivity, Parameter Sharing, pooling and the use of many layers. The basic architecture contains

Figure 3

Basic Architecture of Convolution Neural Network

three layers 1) Convolution layer 2) pooling layer and 3) fully connected layer. Convolution layer have three properties local connectivity, spatial arrangement and parameter sharing: neurons in convolution layer are connected to sub-region of the input image called as receptive field called local connectivity of network for lesser

computations, spatial arrangement is for how neurons arrange in output of convolution layer depends on three parameters which are depth (number of filters used or number of feature maps in first layer of output), second if spatial size which we can find by the equation (w-f+2p)/s +1 where w is input image size, f is filter size or receptive field size as both are same, p is zero padding for pad zero to the border of an image to extract low level features and s is stride which controls how the filter convolves around the input image pixel. After that values founded by convolution layer passed through ReLU (rectified linear unit) to provide non-linearity to a system. second layer is pooling layer (downsampling layer) which is for to overcome overfitting, weight is reduced by 75% and thus lessening the computation cost.max pooling, average pooling is used with same stride and filter (normally 2×2 size). Last layer is fully connected layer which can be multilayer Perceptron and network is trained by backpropagation algorithm.

2.3 Recurrent Neural Network

A recurrent neural network (RNN) is one of the artificial neural networks which form the directed cycle or called feedback Neural Network [13]. RNN is designed to utilize sequential information [10]. RNN is not deep by means of multiple non-linear layers between input and output layer but it is deep by the time stamps. RNN performs computations on both the present and resent past inputs. Which are join to decide how they react to new information, RNNs is that they have a "memory" which stores the data about what has been figured so far [8].

Figure 4

Recurrent Neural Network working

Figure shows how RNN looks after unfolded into a full network or it says unfold network with time for the complete sequence of data. The equations that administer the calculation occurring in a RNN as St = f (Uxt

+WSt-1) [8], St is hidden state at time t, U is weights between input and hidden layer, W is weights between previous and current hidden layers, V is weights between input and output layer and f is hidden layer function tanh or ReLU. RNN shares the same parameters values (U, W, V) at each layers. RNN is train by BackPropagation Through Time (BPTT) algorithm. RNN is used in Language Modeling, Machine Translation, and Image Captioning extra.

3. Applications of Deep Learning

There are several works done which shows us the success of Deep Learning in the various application domains [8]. Applications areas like natural image classification, object detection, biological data classification like biological X-ray, MRI images, biological signals, drug detection, genomes expression recognition, robotics, speech recognition and detection, natural Language Processing, Finger Joint Detection from Radiographs[8], Facial Expression Recognition [1][4].All three Deep Learning Architectures are useful in different application areas for example, Convolution Neural Network is best convenient to classify images because for the image classification it gives highest accuracy [12], like wise Recurrent Neural Network is best suitable for Natural Language Processing, Language modeling and text generating, Machine translation, Speech recognition [13-16] because this network deal with the information in sequence. Human-like Facial Expression Imitation for Humanoid Robot based on Recurrent Neural Network [5].

4. Tools and Technology

Deep Learning Architecture can be implemented in languages like Java, C, C++, MATLAB, and Octave extra. For deep learning algorithms bundle of libraries available this is from different programming languages. Some of libraries are Caffe, DeepLearning4j, Tensorflow, Theano, Keras and Torch. Theano is a Python library that allows you to define, optimize, and evaluate mathematical expressions involving multi-dimensional

arrays which is called Tensors, efficiently overcome the limitations of Numpy. Keras is library written in python which is used to get faster experimentations for deep Networks. Caffe [6] is a deep

learning framework in view of expression, speed, and measured quality. It is created by the Berkeley Vision and Learning Center (BVLC) and by group supporters. Google's DeepDream depends on Caffe Framework. This framework is a BSD-licensed C++ library with Python Interface.

5. Conclusion

The literature on deep learning is vast, mostly coming from the machine learning community. Throughout this review, the important message is conveyed that building and learning deep hierarchies of features are highly desirable. Deep Learning algorithms extract high-level, complex abstractions as good data representations through a hierarchical layers learning process. Deep architectures have outperformed hand- crafted feature engineering in many domains, and made learning possible in domains where engineered features were lacking entirely. CNN gives best accuracy rate to classify images whereas RNN is best suitable for the application of language modelling and generating text, machine translation, Speech recognition, generating image descriptions.

References

Anil, J., and L. Padma Suresh. "Literature survey on face and face ex- pression recognition." Circuit, Power and Computing Technologies (ICCPCT), 2016 International Conference on. IEEE, 2016.

Arel, Itamar, Derek C. Rose, and Thomas P. Karnowski. "Deep ma- chine learning-a new frontier in artificial intelligence research [re- search frontier]." IEEE Computational Intelligence Magazine 5.4 (2010): 13-18.

Bengio, Yoshua, Aaron Courville, and Pascal Vincent. "Representa- tion learning: A review and new perspectives." IEEE transactions on pattern analysis and machine intelligence 35.8 (2013): 1798-1828.

Hsieh CC, Hsih MH, Jiang MK, Cheng YM, Liang EH. Effective se- mantic features for facial expressions recognition using svm. Multi- media Tools and Applications. 2016 Jun 1;75(11):6663-82.

Huang, Zhong, Fuji Ren, and Yanwei Bao. "Human-like facial ex- pression imitation for humanoid robot based on recurrent neural network." Advanced Robotics and Mechatronics (ICARM), International Conference on. IEEE, 2016.

[6] Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Gua- darrama S, Darrell T. Caffe: Convolutional architecture for fast fea- ture embedding. InProceedings of the 22nd ACM international con- ference on Multimedia 2014 Nov 3 (pp. 675-678). ACM.

LeCun, Yann, and M. Ranzato. "Deep learning tutorial." Tutorials in International Conference on Machine Learning (ICML'13). 2013.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learn- ing." Nature 521.7553 (2015): 436-444.

Lee, Sungmin, Minsuk Choi, Hyun-soo Choi, Moon Seok Park, and Sungroh Yoon. "FingerNet: Deep learning-based robust finger joint detection from radiographs." In Biomedical Circuits and Systems Con- ference (BioCAS), 2015 IEEE, pp. 1-4. IEEE, 2015.

Min, Seonwoo, Byunghan Lee, and Sungroh Yoon. "Deep Learning in Bioinformatics."

arXiv preprint arXiv:1603.06430 (2016)..

Mo, Dandan. "A survey o