International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-7 Issue-6S2, April 2019

Artificial Way of Characterizing unsupervised Data using Auto-Encoders With Deep Learning Cluster Analysis

E.Laxmi Lydia, Gogineni Hima Bindu, Pasam Prudhvi Kiran, Kollati Vijaya Kumar

Abstract: ***Most data processing methods for structural, unstructural and semi-structural data are not usually trained to process Big Data. In this 21st century, processing techniques for big data have reached the advanced level of processing data through deep neural networks, which are highly sophisticated in achieving an optimized solution.Autoencoder is a dynamic approach which also combines both supervised learning and unsupervised clustering with minimum reconstruction error.This paper advances the pattern clustering and multidimensional visualization of data.Deep Convolutional Auto-encoder, CDNN-based deep clustering algorithmscomprise of multilayer perceptions improves robustness using Deep Convolutional Embedding Clustering(DCEC), Clustering Convolutional Neural Network (CCNN) clustering algorithms. The objective of this paper is to reduce the computational complexity, enhance reliability, and effective simultaneous feature learning for non-linear transformational data usingautencoders in convolutional networks.***

Keywords: Autoencoders, Convolutional Neural Network,Deep Learning clustering algorithms, MultilayerPerceptron.

I. INTRODUCTION

Pattern clustering consists of various representations of patterns for extraction and selection, defining pattern proximity measure, clustering them and assessment of the data. A traditional approach towork outclustering complex problems usually follows mathematical formulas. Engineering problems cannot be described by equations by adopting adaptive learning architectures such as ANN, SVM, and ELM. Complex problems adopt higher enhanced systems through deep learning. These deep learning algorithms are more demonstrative with compactable architectures 100 times more powerful than ordinary learning algorithms specifically MLP. Multi-Layer Perceptron has the capacity to train algorithms by indulging new neural network architectures. It introduces additional connections beyond layers and maintains different approaches to accomplish complete success than partial success.

Revised Manuscript Received on March 26, 2019

Dr. E. Laxmi Lydia, Department of Computer Science Engineering, Vignan's Institute of InformationTechnology(Autonomous), Visakhapatnam, India.

Gogineni Hima Bindu, Asst. Prof., MCA, Vignan's Institute of InformationTechnology(Autonomous), Visakhapatnam, India.

Pasam Prudhvi Kiran, Asst. Prof., IT, Vignan's Institute of InformationTechnology(Autonomous), Visakhapatnam.

Dr. Kollati Vijaya Kumar, Assoc.Prof, CSE, Vignan's Institute of Engineering for Women, Visakhapatnam, India.

It diminishes gradient problems while training neural networks. It trains without using computations of backpropagation. It describes the size limitation of the algorithm based on the breaks in computational problems.

Large datasets with multiple dimensions are difficult to process. Neural networks design process that can be replaced with/through learning. Special data pre-processing and transformations are used for a specific problem to manage a number of hidden layers. Deep learning neural networks follow the architecture, generate a set of deep features manageto cluster. it calculates the network loss of the clustering by measuring non-clustering loss, clustering-loss and combine both the losses. Later, the Cluster algorithms like CNN, MLP are performed.

Auto-encoder is an active functional unsupervised learning procedure for representation and dimensional reduction.It can increase the rise of progress clustering accuracy by eliminating noisy information that affects the clustering process withproductive time and profitable space complexity**.**

Clustering using a set of deep features

The transformed input is majorly cluster-friendly as the features used for clustering are considered from multiple layers from the network.

Single layer network point to one layer of the network. It is very much beneficial when the dimensions are very low.

Multiple layer network point to the sequence of outputs from all the layers.Accordingly, it allows the entrenched space to define large complex semantic descriptionsthat improve the partitional progress and support similarity computation.

II. BACKGROUND SURVEY

Deep learningclustering algorithms are typicallygrounded into two grades (i) after the specification of learned clustering data, atwo-stage work is applied and (ii) algorithms in deep learning jointly optimize both the feature learning as well as clustering. The anterior grades of algorithms straightly hold the benefits of existing unsupervised deep learning frameworks and different algorithmic approaches.

The usage of autoencoder to determine the lower dimensional features forexisting modeled graph, and then lead by the k-means algorithm to obtain clustering outcomes.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Chen et al, [5] has trained the data layer-wise using DeepBelief Network and clustering the non-parametric maximum-margin to identify the learned intermediary representation. In 2016, Peng et al., demonstrated the usage of autoencoder for the insufficient prior data, while representing nonlinear latent space flexible to both local and global subspace design. At the same time clustering algorithms are occupied to know positioned classlabels. The advanced algorithm approximately a transformed category of Deep Embedded Clustering with a combination of autoencoder to sustain local structure. Yang clearly defined that Deep embedded Clustering has exceeded the terms of clustering accuracy and features representativeness without any recurrent.

A. Krizhevsky et al[1], has suggested that convolutional neural networkshave a biological process with connected patterns among all the neurons. It is extensivelyenforced to image datasets when circumstances and extraction of features (x) shift-variance are needed. They are prepared with a limited clustering loss possiblyensuingallnecessity on loaded, that would automaticallyimprove the clustering accomplishment

B.M.Wilamowski et al[2], they explained several techniques processing the network architectures with effective training by second-orderlearning algorithm and developed a new algorithm for visualization of involving several dimensions of data and also to find clusters that have complex shapes using heterogeneous benchmarks.

Data transformation techniques add linear transformation choose PCA and nonlinear transformation relate to kernel approaches and spectral approaches.Anextremelycomposite latent structure of data is howeverimposed the validity of current clustering mechanisms[3]. Deep neural networks can be practiced to change the data into new clustering-friendly likenessappropriate to its natural intrinsic property of extremely non-linear transformation. Therefore these clustering approaches with deep learning are named after deep clustering[4].

E. Min et al., [7], explained the scientific classification of deep learning algorithms with different network architectures with its characteristics to learn a cluster-friendly representation. Deep Belief networks are vital graphical versions which learn to draw out a deep hierarchical description of the input data. All the guidelines of Deep Belief Networkare calibrated corresponding to fixed loss function.

Elie Aljalbout et al.[14], have described the non-clustering loss with separated clustering algorithms with no further non-clustering loss function. Autoencoder reconstruction loss that contains two divisions encoder and decoder. It is much beneficial when it deals with random noise existence. Self-augmentation loss and clustering losses like k-means loss where data points are delivered to the cluster center through the neural network.

III. METHODOLOGY

Work Process of Autoencoders

Autoencoders are designed based on unsupervised learning approach as they don't need accurate labels to train on. But they develop their own labels for training. The autoencoders require an encoder, code and an decoder. Together the encoder and decoder are linked to the neural network architectures. Visulaization of a autoencoder is given in the following steps:.

Step1: The input data is given to the encoder whichis associated to the neural network architecture to generate the code.

Step2: The decoder, thatequivalent to the neural network architecture, latercreates the output with the help of code.

Step3: Check whether the dimensionality of the input and output are same .

Step4: Finally, the obtained the output equal with the input. There are four parameters to be considered for the framework of autoencoders. They are code size, number of layers in the network, number of nodes per layer and the loss function

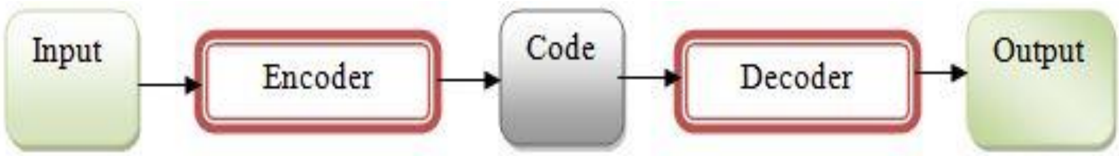

Fig. 1 Functions of Auto-encoders

Work Process of Neural Network architectures

Deep Neural Network involvesvarious clustering approaches with distinct architectures for each method like.

Multilayer perceptron:It containsmultiple layers of neurons in the network. Each layer is processed through the previous layer output irrespective of the first layer. Therefore, it is named as feedforward network.

Convolutional Neural Networkwas influenced by biological terms called neurons, the visual cortex of the animals examined by the higher organizations. It is more practical in regular-grid data for various applications that maintain images by estimating desired features through shift-invariance/ equivariance.

Deep Belief Networkis an unconventional graphical approach**,** maintains multiple layers of suppressed layers. It collects many shallow networks essentially, restricted Boltzmann machines, every sub-network serves as a visible layer to the next sub-network.

Variational Autoencoder is a Bayesian network contains autoencoder architecture that determines the distribution of data.

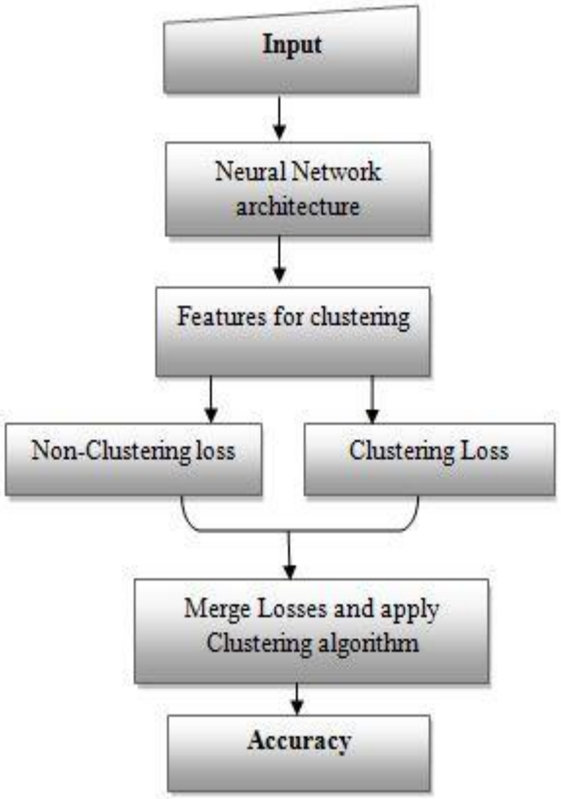

Fig. 2 Flowchart for deep learningclustering-based methods

The above figure 1 describes the workflow of the Deep Learning clustering techniques. Initially, input data is given to the neural network algorithm, by the output of the algorithm, features are extracted using clustering. Clustering Loss, as well as Non-clustering Loss, is calculated and combined to achieve the best accuracy.

Following are the various network architectures carried in the workflow

CDNN-based deep clustering algorithms

CDNN-based deep clustering algorithms mainly work with clustering lossfor cluster assignment hardening loss,nonparametricmaximal margin clustering loss, k-means,equalized information maximization, self-augmented training loss, agglomerative clustering toprepare the networksusing Fully Convolutional Network, Convolutional Neural Network, Deep Belief Network architectures with no network loss. CDNN implements the following algorithms

- **Deep Nonparametric Clustering-** It is an unsupervised pre-trained network improves the classical Nonparametric Maximum Margin Clusteringby initially training a Deep Belief Network into embedding codes to achieve clusters.
- **Deep Embedded Clustering-** It is an unsupervised pre-trained network well-known for deep clustering method, uses autoencoder to reconstruct the cluster loss. The encoded network functions cluster module by considering it as input.It attainsacceptable solutions anddevelops a reference to measure the various new deep clustering algorithm performances.
- **Discriminatively Boosted Clustering-** It is also an unsupervised pre-trained network, pre-trains the auto encoder and operates on cluster assignment and progress more than Deep Embedding Clustering using Convolutional Neural Network.
- **Clustering Convolutional Neural Network-** It is a supervised pre-trainednetwork very much challenging to obtain feasible features from the data. It improves the CNN with permissive large-scale image datasets that makean overall performance to be computationally effective.
- **Information Maximizing Self-Augmented Training-** It is a non-pre-trainednetwork used to achieve function mapping and merges both fully convolutional network as well as RIM to learn and extract cluster assignments. This will suggest self-augment training to establish particular data representations indeep clustering.
- **Joint Unsupervised Learning-** It is a non-pre-trained network works on both images and feature representations. It uses hierarchical clustering. Here two clusters are combined through a predefined loss metric by optimizing clusters with have high computational and memory cost.
- **Deep Adaptive Clustering-** It is a non-pre-trained network well-createdclustering loss, accomplish the state of art completion on assorteddatasets.

Deep Convolutional Embedding Clustering (DCEC)

Deep Embedded Clusteringbegins with pre-training an autoencoder and then eliminates the decoder. The existing encoder is calibrated by computing the successiveintention:

$$L = KL(P||Q) = \sum_i \sum_j p_{ij}\, log \frac{p_{ij}}{q_{ij}}$$

q_{ij} describes the connectivityamongz_i(embedded point)and μj (cluster center) determined by t-distribution i.e, $q_{i,j} = \frac{\left(1+\left|\left|z_i-\mu_j\right|\right|^2\right)^{-1}}{\sum_j\left(1+\left|\left|z_i-\mu_j\right|\right|^2\right)^{-1}}$

the target distribution is evaluated by pi,j in the autoencoder equation, $p_{i,j}= \frac{{q_{i,j}}^2 / \sum_i q_{i,j}}{\sum_j({q_{i,j}}^2 / \sum_i q_{i,j})}$

Decreasing L in the target distribution P is a form of self-training. Suppose if encoder mapping $f_w$and x_i be the input then embedded score z_i= $f_w(x_i)$, where X contains the dataset elements. Pre-training of every element in z_i can also be obtained by using features f_w . Later, k-means is applied to embedded elements$\{z_i\}$ to obtain primary clusters$\{\mu j\}$. In the process of backpropagation, $\frac{\partial L}{\partial zi}$ and $\frac{\partial L}{\partial \mu j}$ can be quicklymeasured**.** Here $f_w$is updated using $\boldsymbol{\frac{\partial L}{\partial zi}}$ when it is passed down and cluster center$\mu j$is updated using$\frac{\partial L}{\partial \mu j}$i.e, $\mu_j = \mu_j - \lambda \frac{\partial L}{\partial \mu_j}$

An important input of DECalgorithm is related to the clustering loss and target distribution.It is functioned by using extremely confidential samplingfor managing and creating the samples for each cluster allocation. Nevertheless, samples are not drawn closer to the margins close to the cluster center. The problem here is maintaining the local structure of the data, due to this the marginal samples to get closer to the appropriate cluster.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Clustering Convolutional Neural Network (CCNN) algorithm

Clustering Convolutional Neural Network initializesthe layersbased on k-means centroids on convolutional neural network softmax output layer. The output layer of the network within are used as features. This network utilizes the feature extractors to enhance clustering by implementing hierarchical techniques and agglomerative clusterin loss function.

Non-Clustering Lossis liberative for clustering algorithms and accomplishes wanted constraint on the experienced method. Following are the some of the potential options carried for non-clustering loss

- **No non-clustering loss** is a loss function for N additional non-clustering loss. The clustering loss from the network is constrained. Dealing with clustering loss, the dearth of non-clustering loss has lowest and unfavorable results or apparently much weaken the clusters(Yang et al., 2016a [5]), but in practice, modern one is carried frequently.
- **Autoencoder reconstruction loss:** autoencoder mainly focus on two functions. They are encoder and decoder. The encoder arranges the input x to latent space Z by using z as representation. While training the data, the decoder will attain During training, the decoder tries to modernize x from z, this will strongly make confirmation of information lost through the encoding phase. Once the data is been trained the decoder will not be used anymore and the encoder will map the input to Z. Performing this approach, the autoencoders will achieve beneficial representations when arbitrary noise is added to the input.

Moreover, autoencoders manage to overcome the problems of dimensionality reduction goals.Typically, the reconstruction lossis estimated by the distance measure defined as, $\boldsymbol{d_{AE}(x_i, f(x_i))}$in the midst of input x_i to the autoencoderincluding the equivalent reconstruction function $f(x_i)$.Mean squared error is the precise formulation for these two variables:

$$L= d_{AE} (x_i, f(x_i)) =\sum_i \left|\left|x_i - f(x_i)\right|\right|^2$$

The determined loss function ensures that the accomplished representation extract essential data through the introductory one, thus due to this reconstruction is attainable.

Clustering Loss is a standard function describes clustering approaches and available representations . some of the clustering loss functions are No clustering loss, k-means loss, cluster assignment hardening, Balanced assignments loss, Locality-preserving loss, Group sparsity loss, cluster classification loss, Agglomerative clustering loss.

Mechanism to combine Losses clustering loss ($L_c(\theta)$) and non-clustering loss ($L_n(\theta)$) function are associated using $L(\theta)= \alpha L_c(\theta) + (1-\alpha)L_n(\theta)$

Where $\propto \in [0; 1]$ is a stable weighting among the loss functions. This can also be replaced while training. The values of $\propto$the canare scheduled while pre-training, fine-tuning, joint training, variable schedule.

Even after the clustering progress, we again apply clustering algorithm i.e, re-run the process from pulled using learned features.

→ Clustering identical dataset using mapping.
→ Acquiring better outcomes using a number of iterations.

IV. SOFTWAREENVIRONMENT

In the development and implementation of the algorithms, a number of different software have been used. This work is carried out in Tensor flow and Keras, which areGoogle-developed open source librarys for machine learning. The underlying idea is to think of it as dataflow graphs where the nodes correspond to numerical operations and the edges, called tensors, corresponds to multidimensional arrays connecting the nodes.The library can be used in the programming languages keras, Python, C and C++.

V. CONCLUSION

Deep Clustering is broadlyimplemented in major experimentaloperations for its vigorous intelligence in extracting features, it is logical and instinctive to integrate clustering algorithms upon deep learning for more excellent clustering results. In this paper, we provide a standardized analysis ofautoencoders usingdeep clustering the pattern clustering is getting advanced. The Clustering of non-linear transformational data specificallypresents the typical features and dominance. Current Deep Clustering Algorithms often target on image datasets, at the same time few experiments have been shaped on continuous data like documents. To this impact, it is supported to examine the utility of connectingalternative more network architectures throughDeep Clustering. Thus, optimizing networks and clustering techniquesapproximatelysustain the clustering performance.

REFERENCES

1. A. Krizhevsky, I. Sutskever, and G.E. Hinton, “ImageNet Classification with deep convolutional neural networks” in Proc. Adv. Neural Inf. Process. Syst., 2012, pp.1097-1105
2. B.M. Wilamowski, Bo Wu, Janusz Korniak, “Big Data and Deep learning”, INES 2016, IEEE International Conference on Intelligent Engineering Systems, 2016.
3. J.R. Hershey, Z. Chen, J. Le Roux, and S. Watanabe, “Deep clustering: Discriminative embeddings for segmentation and separation ” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Mar. 2016, pp.31-35
4. J. Schmidhuber, “Deep Learning in neural networks: An overview”, Neural Network., vol.61, pp.85-117, Jan 2015.
5. D. Chen, J. Lv, and Z. Yi, “Unsupervised multi-manifold clustering by learning deep representation”, in Proc. Workshops 31st AAAI Conf. Artif.Intell. (AAAI), 2017, pp.385-391.
6. C.C. Hsu and C. W. Lin, “CNN-based joint clustering and representation learning with feature drift compensation for large-scale image data,” IEEE Trans. Multimedia, vol.20, no.2,pp. 421-429, Feb.2018
7. Erxue min, Xifeng Guo Qiang Liu, Gen Zhang, Jianjing Cui, and Jun Long, “A Survey of Clustering with Deep Learning: From the perspective of Network Architecture.”, translations and content mining volume 6, 2018
8. B.Yang, X. Fu, N.D. Sidiropoulos, and M.Hong.(2016), “Towards K-means-friendly spaces: Simultaneous deep learning and clustering”, available: https://arxiv.org/abs/1610.04794
9. E.Aljalbout, V. Golkov, Y. Siddiqui, and D. Cremers, (2018), “clustering deep learning: taxonomy and new methods”.

Available:https://arxiv.org/abs/1801.07648

10. V.IIango, Dr. R. Subramanian, Dr.V.Vasudevan, "Cluster Analysis Research Design model, problems, issues, challenges, trends and tools", International Journal on Computer Science and Engineering (IJCSE),
11. J. Xie, R. Girshick, and A. Farhadi, " Unsupervised deep embedding for clustering analysis", in Proc. Int. Conf. Mach. Learn., 2016, pp.478-487.
12. Dizaji, K.G., Herandi, A., Deng, C., Cai, W., and Huang, H.(2017). Deep clustering via joint convolutional autoencoder embedding and relativeentropy minimization. In 2017 IEEE International conference on computer vision (ICCV), pages 5747-5756. IEEE.
13. K. Simonyan and A. Zisserman. (2014). "Very deep convolutional networks for large-scale image recognition". https:arxiv.org/abs/1409.1556
14. Elie Alijalbout, Vladimir Golov, Yawar Siddiqui, Maximilian Strobel & Daniel Cremers, "Clustering with Deep Learning: Taxonomy and New Methods" arXiv:1801.07648v2[cs.LG] September 2018.
15. Chunfeng Song, Feng Liu, Yongzhen Huang, Liang Wang, and Tieniu Tan, " Auto-Encoder Based Data Clustering", J. Ruiz-Shulcloper and G. Sanniti di Baja(Eds.): CIARP 2013, part I, LNCS 8258, pp. 117-124, 2013, Springer- Verlag Berlin Heidelberg 2013.
16. Chen, D., Lv. J., and Yi, Z.,(2017), Unsupervised multi-manifold clustering by learning deep representation. In workshops at the AAAI Conference on Artificial Intelligence, Pages 385-391.
17. Song, C., Liu, F., Huang, Y., Wang, L., and Tan, T. (2013), Autoencoder based data clustering. In Iberoamerican Congress on Pattern Recognition(CIARP), Pages 117-124, Springer.
18. Xie, J., Girshick, R., and Farhadi, A.(2016). Unsupervised deep embedding for clustering analysis. In an international conference on machine learning (ICML), pages 478-487.
19. J. Guerin, O. Gibaru, S. Thiery, and E. Nyiri, (2017). "CNN features are also features are also great at unsupervised classification", https://arxiv.org/abs/1707.01700
20. B. Wu and B. M. Wilamowski, " An Algorithm for visualization of Big Data in a two-dimensional space." The 41st annual conference of the IEEE IECON 2015, pp.53-58, 9-12 Nov 2015
21. X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks", inan international conference on artificial intelligence and satisfies,2010, pp.249-256.