Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 4 (2023)

Adversarial Regularized Class Incremental Autoencoder Technique with Convolution Neural Network for Secure Clustering of Short Text in the High Dimensional Data

[1]Kiruthika B, [2]Dr. B. Srinivasan, [3]Dr. P. Prabhusundhar

[1] [3]Assistant Professor in Computer Science
[2]Associate Professor in Computer Science
Gobi Arts & Science College (Autonomous), Gobichettipalayam, India.

E-mail: [1]kiruthikabalu@gmail.com [2]srinivasanb@gascgobi.ac.in
[3]drprabhusundhar@gascgobi.ac.in

Abstract: High Dimensional data Clustering with differential privacy has been gained significant attention recently in the large scale distributed cloud data center to cluster the high dimensional data with increased security to the data outsourced. Especially deep learning architecture based on deep adversarial regularized hierarchical autoencoder structure provides efficient clustering solution to high dimensional data containing non linear long text. However it provides class imbalance problem in terms of over-fitting and under-fitting issues with high computation complexity. To mitigate those challenges, class incremental deep adversarial regularized multi-view autoencoder technique is proposed in this work. Initially, missing value prediction using factor analysis and dimensionality reduction and data normalization is carried out to high dimensional data using principle component analysis. Principle component analysis technique eliminates the reconstruction errors occurring due to short text in the dimension space to project the sparse matrix. Sparse matrix is used to the select the discriminative features using non linear discriminant analysis. Selected feature is employed to the convolution neural network to process the selected features to establish the cluster using fully connected layer with activation function and softmax layer. Convolution Neural Network is capable of clustering structurally similar attributes and its information containing the normalized short text with increased classes. Proposed model achieves minimized intra cluster similarity and inter cluster variation by computing the data affinity of new representation. Further autoencoder model is to encode cluster structure containing attribute information by reconstructing the adjacency matrix. Decoder model retrieves or suggest the data records on basis of the attribute similarity or structure similarity of the record attribute or user attribute of the high dimensional data. Detailed experimental analysis has been performed on benchmarks datasets to compute the proposed model performance with conventional approaches using cross fold validation. The performance outcome represents that proposed architecture can produce good accuracy and effectivenesson high dimensional data containing the short-text.

Keywords: High Dimensional Data Clustering, Big Data Cloud, Privacy Preserving, Variational Autoencoder, Advanced Networking

1. Introduction

Emergence of evolving high dimensional data in various area has brought several challenges to the existing machine learning and deep learning based clustering approach[1]. Recently considerable efforts has been made in research paradigm on basis of the multiview clustering to handle diverse and inconsistent data. Inconsistent data represented as evolving short text[2]. Multiview clustering is highly capable of learning local structural affinities among the data points. Beside the emphasize of the multiview clustering to the inconsistent data rather limited attention has been paid to data privacy in the distributed big data cloud environment which leads to data disclosure attacks[3].

Traditionally secure clustering techniques[4] based on deep adversarial regularized hierarchical autoencoder structure has been developed and implemented to the large distributed cloud environment to cluster the high dimensional data with increased security to the outsourced data against data disclosure attacks occurring on processing extensive transaction queries to the cloud server. Despite of its numerous advantageous, still it becomes complex in handling the short text in data points and data queries to the outsourced non linear high dimensional data[5]. Further it produces class imbalance problem in terms of over-fitting and under-fitting issues with high computation complexity.

Class incremental deep adversarial regularized multi-view autoencoder technique is proposed in this work which is capable of handling the short text clustering and securing the outsourced data. Initially, missing value prediction using factor analysis[6] and dimensionality reduction and data normalization using principle component analysis[7] is carried out to preprocess the high dimensional data. Preprocessed data in form sparse matrix is employed to non linear discriminant analysis[8] to select the discriminative features.

Selected feature is projected to the convolution neural network to establish the cluster on processing the structurally similar attributes and its information in the fully connected layer with encomposing the activation function, loss layer and softmax layer. Finally autoencoder approach is integrated into existing architecture to encode cluster structure containing attribute information by reconstructing the adjacency matrix. Decoder model retrieves or suggest the data records on basis of the attribute similarity or structure similarity of the record attribute or user attribute of the high dimensionaldata.

2. Related work

In this section, high dimensional data clustering models in the big data cloud environment using deep learning and machine learning methods has been analyzed in detail on its generated discriminate clusters along the processing of the data along the preprocessing and feature extraction process to generate the efficient clusters. Technique which performs similar to proposed model for high dimensional data clustering is described as follows

2.1. Deep Privacy Preserving Convolutional Autoencoder Learning Technique Along Anonymization for Outsourcing High Dimensional Data

In this literature**,** deep privacy preserving convolutional auto encoder learning model will process the data to establish the secure high dimensional data cluster on inferring the data distribution extracted from the cloud environment[9]. Initially privacy preserving of the confidential information containing in the high dimensional space is secured using anonymization approach. Anonymized secure attribute is subjected to feature learning along autoencoder for high cluster friendly illustrations on generation of objective function to produce maximum margin cluster. Those cluster further fine tuned using stochastic gradient descent towards feature refinement on the hyperparameter of various layers of deep learning model to establish the minimum reconstruction error by feature refinement[8]. It further minimizes the intra cluster correlation and inter cluster covariance in the feature space for cluster assignment..

2.2. Deep Adversarial Regularized Autoencoder Technique

In this literature, deep adversarial regularized autoencoder technique explore the latent data structure and computes the associations of the data points to construct the spatial and temporal cluster structures to high dimensional data[10]. To approximate cluster structures, hyperparameter tuning using RMSProp is carried out in the output layer to achieve the data instance in the cluster to be nearer to each other by computing the data affinity on new representation. Cluster results produce the minimized intra cluster correlation and inter cluster covariance in the feature space. In particular, autoencoder technique ishighly efficient in securing the processing data.

3. Proposed Model

In this section, architecture of proposed class incremental deep adversarial regularized multi-view autoencoder technique for high dimensional data clustering on inclusion of convolution neural network for soft text clustering and securing the clustered structure using the variational encoder and decoder layers to preserve the outsourced information on the big data cloud environment has been illustrated in thefigure 1 as

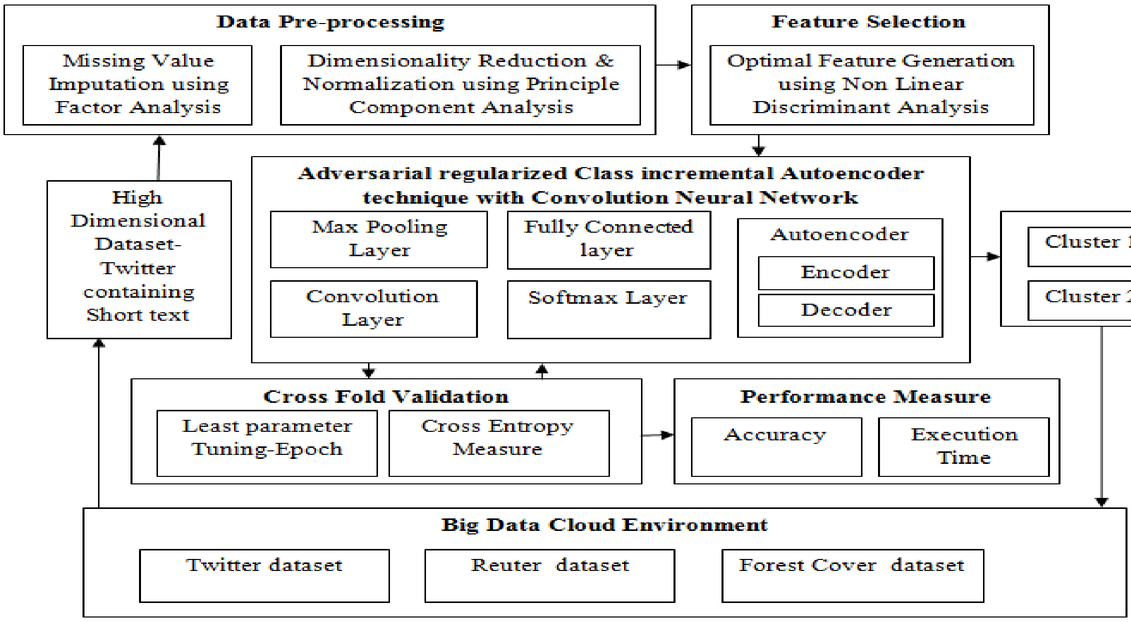

Fig 1: Architecture of the proposed methodology

3.1. Data Pre-processing

Big Data cloud architecture is capable of handling the high dimensional data which contains both irrelevant attribute and noise attributes with missing value. Missing value prediction using factor analysis[11] as it predicts the data for missing field..

- **Factor Analysis for Missing Value Imputation**

Factor analysis is projected in work for Missing Value Imputation in the irrelevant and relevant attributes. Factor Analysis computes maximum common variance on the specified data field. It employs the Kaiser criterion to the Eigen value of the data matrix. Further it estimates the value to the missed data point of the specified attribute through mean and variance[12]. It computes maximum probability value on processing the matrix towards data correlation to the missing data field.

3.2. Dimensionality Reduction and Data Normalization

Dimensionality reduction is employed to eliminate the curse of dimensionality issues and data normalization is to eliminate the reconstruction errors occurring due to short text in the dimension space. In this work, principle component analysis is employed to achieve both dimensionality reduction and datanormalization in parallel.

3.2.1. Data Normalization

Data normalization is carried out to high dimensional data using PCA. Normalization will eliminate the reconstruction error in the feature space and it standardizes the dataset to measure the entire feature in the same scale. Initially mean of the feature has to be computed and then it has to be subtracted from every data point of the particular feature. Next, every data point has to be divided to standard deviation[13].

$$\text{Minimum Error Objective } ME(x)= \frac{1}{N}\sum_{e=1}^{N} \|x_e - \tilde{x}_e\|^2 \cdots\cdots\cdots\cdots \text{Eq.1}$$

where x is the reconstruction generated from the mean and standard deviation of the data points.

Resultant dataset will be linear to each other after mean and standard deviation computation on setting the derivates zero is considered as constraint to the each feature containing various data points.

3.2.2. Dimensionality Reduction

Principle Component Analysis is employed to minimize the high dimensional dataset irrelevant and noise attribute into relevant attribute set. It is carried out on computation of maximum variance in the high dimensional data on construction of transformation matrix[14]. Matrix generates the vector containing maximum variance features or attributes which considered as new subspace.

Original high dimensional feature space is represented as x= {x1,x2,........xd}, $x \varepsilon R^d$ Eq.2

Reduced feature space is represented as z={z1,z2....zk} $z \varepsilon R^k$ Eq.3

In reduced feature space, principle components derived as selected attributes will have largest possible variance. Further all consequent principle component of the feature space will have largest variance along the constraint that principle component selected should be uncorrelated to other principle component. Finally it becomes mandatory to standardize the features as entire feature of the highdimensional data should be measured in same scale.

Covariance Matrix for d number of dimension in dataset stores the pair wise covariance between different features is as follows

Covariance matrix C_m of two features x_j and x_k $= \frac{1}{n}\sum_{i=1}^{n}(x_j^i - \mu_j)(x_k^i - \mu_k)$................Eq.4

where μ_j *and* μ_k are mean of the feature j and k respectively. A positive covariance among two features represent that the features mean enhances or minimizes together, whereas a negative covariance represents that mean of the features change in other directions of the maximum variance.

Next, decompose the covariance matrix into Eigen vector and Eigen value. The eigenvectors of the covariance matrix termed as the principal components (the directions of maximum variance), whereas the corresponding eigenvalues will represent their magnitude. After computation of Eigen value for Eigen vector, eigenpairs of the covariance matrix has to be obtained . It is obtained on satisfying the following condition

$$\sum V = \lambda V \qquad \text{Eq.5}$$

where λ is scalar which is represented as Eigen value.

Eigen values of the eigen pairs are sorted by decreasing magnitude. In that top k eigen vectors is selected on basis of the eigen values and it is considered as informative attribute. Finally projection matrix is constructed on the selected Eigen vector, it is considered as dimensionality reduced dataset (k dimensional feature space) for the clustering task of the dataset.

3.3. Non Linear Discriminant Analysis

Non linear discriminant analysis is subset of the original feature which retains the most relevant information. Sparse matrix is used to the select the discriminative features using non linear discriminant analysis[15]. Non linear discriminant analysis uses scatter matrix to compute the relationship between the Eigen vectors. Dimensional mean vector is computed using fisher criterion function on the projected Eigen vector. Vector is generated to maximize the separation between vectors of the projected matrix.

Fisher's criteria maximize the distance of the projected mean of the data points among the vector and minimize the projected mean value of the data points in the vectors.

Scatter matrix of data instances within the attribute (Eigen vector) = $S_W = \sum_{i=1}^{c} S_i$ Eq.6

Scatter matrix of data instances among the attributes (Eigen vector) = $S_B = \sum_{j=1}^{c} N_i (mj-m)(mj-m)^T$..Eq.7

The optimal solution vector containing the attribute is obtained on reduction of the feature space through usage of scatter matrix.

3.4. Convolution Neural Network

Convolution Neural Network is a deep learning architecture employed for clustering the feature selected using non linear discriminant analysis approach. In this part, Convolution Neural Network uses parametric tuning of the activation function towards cluster generation to the short text in the various attributes of the dataset on inclusion of activation functions. and these representations have been processed.

- **Max Pooling Layer**

Max pooling layer has been implemented to maps the highly subjective features of feature selection process into structurally similar features representation containing the similar short text for cluster generation. It enhance the separation of data points during the similarity computation on data clustering inthe fully connected layer[16]. Hyperparameterized components of the Convolution Neural Network for cluster generation is represented in the table 1.

Table 1: HyperParameterized Component for the Convolution Neural Network

Hyper Parameter	Values
Batch Size	148
Model Learning Rate	10^{-4}
Attribute Size	70
No of Epoch	50
Loss function	Cross entropy

- **Convolution Layer**

The Convolution layers gathers the structurally similar attributes and it learns to transform the low level subjective features to more abstract features with reduced hyperparameter. Each attribute or feature is considered as tensor and computes the weight of the feature for the each data point. Filters eliminate the reduced weighted features. Finally outcome of the layer is considered as feature space with depth dimension which has large weight.

Depth Dimension
□ □$u^m d^2$
+ λ(□u
□1) c c
ik ik
i□1
ik
i□1
..................Eq.8

- **Fully Connected Layer**

Fully connected layer is composed of activation function, softmax layer and loss layer to generate the cluster. It extract the deep features of the convolution layer to cluster it in the softmax layer using the activation function. The optimization of hyper parameter used in the fully connected neural network is given by

$C = \gamma Lc + (1-\gamma)Lc$ Eq.9

Where γ is considered as hyperparameter and L_C is considered as class limit.

- **Softmax Layer**

Softmax layer uses the Euclidean distance to cluster the features containing the data points. Data point considered in this work is short text. It is capable of clustering structurally similar attributes and its information containing the normalized short text with increased classes. Proposed model achieves minimized intra cluster similarity and inter cluster variation by computing the data affinity of new representation. Clusters are established with cluster limit. The cluster limit is set on basis of the weight function specific to feature and its bias.

$L_C = \sigma(wF+b)$ Eq.10

Where L_C is cluster limit, σ is the cluster , f is the feature space and b is the bias function to map the features to the cluster on basis of similarity. Cluster generation is carried out using word vector and epoch. Softmax layer minimizes the intra cluster compactness and inter cluster seperability in the feature space. Softmax layer minimizes the reconstruction error on the computing the mean and variance.

- **Activation Function**

The proposed approach uses the Rectified Linear Units (ReLU) activation function, which identifies non-linearity to the features. Feature vector containing the subjective attributes of the dataset has been processed with parameterized values to generate the appropriate cluster to the short text.. Activation function is employed to train the model to generate cluster as it is several times faster than other functions.

- **Output Layer**

The output layer of the convolution neural network contains the data cluster representing the data points of the attributes. Hyperparameter tuning has been enabled in the output layer to make the datainstance in the cluster to be close to each other by computing the affinity of the data on new representation. It results significant increase in the clustering performance on the discriminative information's.

- **Loss Layer**

Loss layer is to ensure the cluster similarity with respect to intra cluster compactness and inter cluster seperability. Further it is to fine tune the hyperparameter on different layers of convolution neural network to ensure the minimum reconstruction error. In this work, cross entropy loss function has been utilized to manage cluster seperability of high dimensional data. Model parameter of neural network is updated to generate the distributed cluster with minimum inter cluster distance for novel data points.

Hyper parameter is important for the balance between the within-cluster distance and the between-cluster distance in the cluster space. It is given by

$$Q = \sum_{i=1}^{c}\sum_{k=1}^{N} c_{ik}^{m} \|\mathbf{x}_k - \mathbf{v}_i\|^2$$

......................Eq.11

Where C is subject to cluster partitions and its prototype

3.5. Autoencoder

Autoencoder is applied to secure the cluster structure generated against data disclosure attacks. Variational Autoencoder model is considered as best transformation vector containing the encoder block, decoder block and latent spaces contain the latent variables. Cluster structures containing attribute information and its data points are encoded using the encoding operation[17]. Attribute of the cluster is reconstructed to latent variable on basis adjacency matrix is given as

Adjacency matrix = Q[L|x,A] =N(L,μ,σ^2 I}.......................Eq.12

Original Input Data X= {Attribute$_1$, Attribute$_2$ .... Attribute}

Latent Space of the data L= {Attribute probability$_1$, Attribute probability$_2$… Attribute probability}Where μ is the mean and σ^2 is the variance and I is the identity matrix.

Encode Operation Z = h($W_{encoding}$ X+ $b_{encoding}$) Eq.13

Where $W_{encoding}$ and $b_{encoding}$ is the parameter and h is the activation function

Encoded cluster is stored in the big data cloud environments. Decoder model retrieves or suggest the data records on basis of the attribute similarity or structure similarity of the record attribute or user attribute of the high dimensional data.

Decode(z) = h($W_{decoding}$Z + $b_{decoding}$)...Eq.14

Where $W_{decoding}$ and $b_{decoding}$ is the parameter and h is the activation function

Probability distribution of the hidden attributes after decoding operation is given by KL divergence by

Probability distribution of attribute vector P(z|x) = $\frac{p(x|z)p(z)}{p(x)}$....Eq.15

The variational attributes selected on inference to approximate the conditional probability distribution of the latent variables. KL divergence determines the disparity between the attributes and it minimizes the difference as best as possible..

Algorithm 1: Deep Non Parametric Learning

Input: High Dimensional Dataset
Output: Secure Data Clusters in Encoded FormProcess
Data Pre- Process () Compute Missing value ()
Assign Kaiser Criterion to Eigen Value of the Data Matrix
Set largest variance of the Eigen value to specified missing field is set as imputation value
Dimensionality Reduction using PCA()
High Dimensional data to low dimensional dataCompute Principle Component()
Covariance(Eigen Vector)Rank Eigen Vector
Select top K eigen Vector as principle ComponentsReduced features = (PE_{v1}, PE_{v2} PE_{vn})
Feature Selection _LDA ()

Scatter matrix M[] = $\frac{1}{N} \sum (\boldsymbol{\delta}_i - \boldsymbol{\delta}_j)(\boldsymbol{\delta}_i - \boldsymbol{\delta}_j)^T$

Compute Covariance of Scatter matrix M[] is the subjective features

Generate subspace for F as subjective featuresApply Convolution Neural Network ()
Max Pooling ()
Compute high level subjective features

Convolution Layer()
Estimate feature using kernel functionSelected feature for clustering is Sf
Fully connected Layer() Activation Layer_ReLu()
Softmax layer()Cluster (Sf)
Compute Inter cluster compactness and intra cluster similarity()Loss Layer_Cross Entophy()
Output Layer() Cluster = {c1,c2}
Privacy Preserving ()
Adjacency matrix (original attributes of the cluster) = Latent AttributesEncode (latent Attributes)
Store cluster containing encoded latent attribute in the data centerDecode (latent Attributes) related to the search
Output the original attribute and its information of the cluster.

4. Experimental results

Experimental analysis of current architecture represented as adversarial regularized class incremental autoencoder technique with convolution neural network to large scale distributed big data cloud environment for secure clustering of short text in the high dimensional data has been formulated using the twitter dataset represented in CSV pattern [18]. Performance of the proposed model has been computed on basis of Precision, Recall and F-Measure using cross fold validation.

Cross fold validation is carried out on the confusion matrix of the validation data. In this work, 60% of twitter dataset is subjected to training and 20% of dataset is subjected to testing and remaining 20% of dataset is used for validation of the current architecture. The training parameter of the privacy preserving deep learning approach has been specified in the table 2

Table 2: Training parameters

Parameter	Value
Model Learning rate	10^{-4}
Loss Function	Cross Entropy
Activation function	ReLu
Epoch	50

4.1. Dataset description: Twitter

High dimensional dataset composed of multiple attributes was collected from Twitter website. The data set contains 340,000 tweets of various health trajectories on wide classes in variety of area and patients on world in the covid -19 pandemic period. Data is represented in CSV format.

4.2. Performance Evaluation

Performance of the current architecture has been evaluated on basis of accuracy of the high dimensional data clustering on reducing the reconstruction error and curse of dimensionality issues. Accuracy of the clustering depends on the inter cluster compactness and intra cluster similarity. Further accuracy is computed using confusion matrix represented using the validation data of the dataset. Confusion matrix outcomes the results in terms of the true positive, true negative, false positive and false negative.

High discrimination among the data points in cluster is eliminated by hyperparameter tuning of various layers of the model. Parameter tuning prevents the overfitting issue and it prevents the under fitting issues in cluster generation on processing the short text value of the dimension in the dataset. In order to preserve the privacy of the transactional queries, rank constraints for high level features is considered. Accuracy of the model is computed on basis of following metric.

- **Precision**

Precision is considered as ratio of number of optimal feature selected for clustering to the number of feature subspace extracted. It is also defined as ratio of relevant data points of the attributes grouped into cluster to the set of the attributes of the entire dataset[19]. In order words, it is illustrated as the fraction of relevant data points of the subjective attributes among the every cluster structures generated. Figure 2 depicts the performance of the current model with respect to precision measure on processing the validation data of the twitter dataset.

$$\text{Precision} = \frac{\text{True positive}}{\text{True positive+False Positive}} \quad \text{.......................Eq.16}$$

True positive is considered as number of related data points of the feature or attribute and false negative is considered as unrelated data points of the feature or attribute of the dataset under processing [15]. Importantly clustering performance is characterized by greater intra-cluster similarity and lower inter-cluster similarity on the data points of the cluster generated.

- **Recall**

Recall is considered as ratio of number of feature unselected for clustering to the number of feature subspace extracted. It is also defined as ratio of irrelevant data points of the attributes to the set of the attributes of the entire dataset. In order words, it is illustrated as the fraction of irrelevant data points of the attributes to the whole attribute of the dataset. Figure 2 depicts the performance of the current model with respect to recall measure on processing the validation data of the twitter dataset.

$$\text{Recall} = \frac{\text{True positive}}{\text{True positive+False negative}} \quad \text{..................Eq.17}$$

On employing the deep learning architecture, efficiency of the data cluster depends on the activation function and loss layer of the model. Further encoder model towards privacy preserving of the data is computed using high level features. High level feature of the data is transformed into the latent variables in the adjacency matrix to create encoded data . F measure is a measure of the cluster quality using epoch and batch size of the cluster. It is uses results of the precision and recall to calculate it. High value of the f-measure represents high cluster quality.

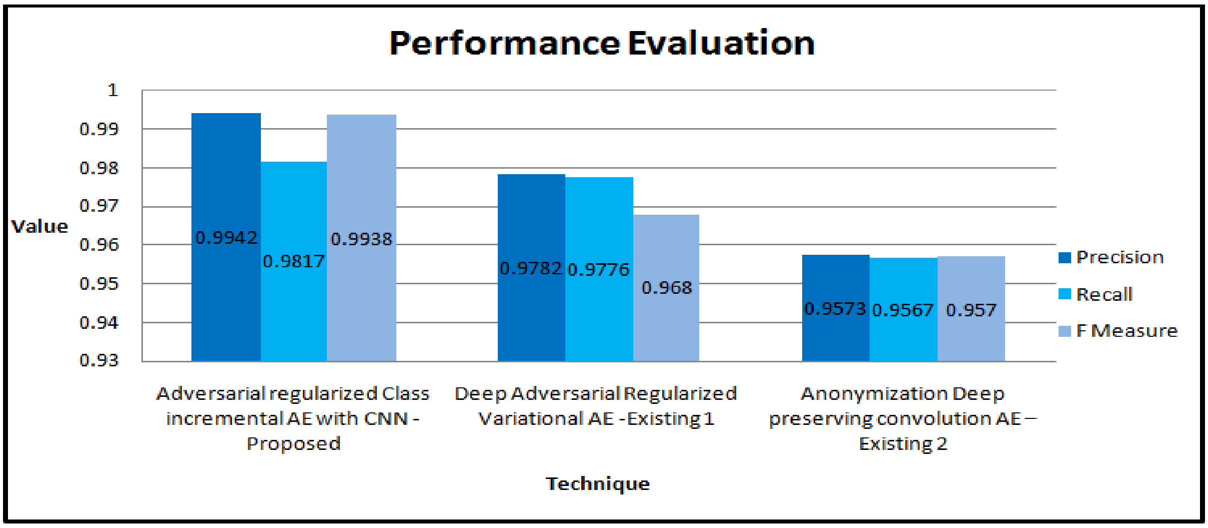

Fig 2: Performance Evaluation of the methodologies

- **F measure**

It is the number of relevant data points in the clusters of entire data in the learning model. It is considered as accuracy. Figure 2 depicts the performance of the current model with respect to recall measure on processing the validation data of the twitter dataset. Accuracy is computed by

$$\text{Accuracy} = \frac{\text{True positive} + \text{True Negative}}{True\ positive + True\ Negative + false\ positive + False\ negative} \quad \text{...Eq.18}$$

Especially, data points in terms of short text will leads to major challenges in cluster formation. However proposed architecture is highly capable of the handling the short text and it is highly efficient in producing the word vector to each attribute or feature for cluster formation. Further, Autoencoder technique reduces the size of the encode data on transformation of the original attribute space of cluster to latent attribute space before encoding. Table 3 represents the performance evaluation of the high dimensional clustering techniques to twitter dataset.

Table 3: Performance Analysis of Adversarial regularized Autoencoder approaches

Technique	**Precision**	**Recall**	**F measure**
Adversarial Regularized Class Incremental Autoencoder Technique with Convolution Neural Network - Proposed	0.9942	0.9817	0.9938
Deep Adversarial Regularized Variational Autoencoder Model - Existing 1	0.9782	0.9776	0.9680
Anonymization Deep Preserving Convolution Autoencoder –Existing 2	0.9573	0.9567	0.9570

On analysis of the current clustering approach on both theoretical and experimental aspect it is high capable in clustering the high dimensional dataset into high representative clusters[20] and in addition, it is capable of securing the high level feature of the data against the data disclosure attacks in the big data cloud environment.

5. Conclusion

Adversarial regularized class incremental autoencoder technique is designed and implemented in this work to generate secure high representative clustering in the big data cloud environment. On processing the current model, it is capable of exploring the deep latent structure of the high dimensional data and computes the associations of the data points to construct the high representative cluster structures to high dimensional data using the principle component analysis technique . Further evolving data streams are approximated using the variational autoencoder to preserve the cluster structures. Approximated data points of the attributes is extracted and selected for optimal subjective feature for classification is achieved using non linear discriminant analysis. Fully connected layer of convolution neural network using activation function and softmax layer with loss function is employed to generate the efficient data clusters with minimized intra cluster similarity and inter cluster variation to the feature space. Further, model is hyper parameter tuned in the output layer to make the data instance in the cluster to be close to each other. Finally proposed model proves that it is effective and high scalable on high dimensional data on achieving the high accuracy in the generated cluster structures.

References

[1] Min E, Guo X, Qiang “A survey of clustering with deep learning: from the perspective of network architecture” in IEEE Access, Vol. 6, issue.39, pp: 501–14, 2018.

[2] Chowdary NS, Prasanna DS, Sudhakar P. “Evaluating and analyzing clusters in data mining using different algorithms”. International Journal of Computer Science and Mobile Computing, Vol.3, PP: 86–99, 2014.

[3] Davidson I, Ravi SS. Agglomerative hierarchical clustering with constraints: theoretical and empirical results. In: European Conference on Principles of Data Mining and Knowledge Discovery. Heidelberg, Germany: Springer, 2005, pp: 59–70.

[4] Dizaji KG, Herandi A, Cheng,” Deep clustering via joint convolutional autoencoder embedding and relative entropy minimization. In: 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017, 5747–56.

[5] Xie J, Girshick R, Farhadi A ”Unsupervised deep embedding for clustering analysis. In: International Conference on Machine Learning. New York City, NY, USA: ICMLR, 2016, 478– 87.

[6] K. Buza, A. Nanopoulos, and L. Schmidt-Thieme, “INSIGHT: Efficient and Effective Instance Selection for Time-Series Classification,” Proc. 15th Pacific-Asia Conf. Knowledge Discovery and Data Mining (PAKDD), Part II, pp. 149-160, 2011.

[7] L. Parsons, E. Haque, and H. Liu, “Subspace clustering for high dimensional data: a review,” ACM SIGKDD Explorations Newsletter, vol. 6, no. 1, pp. 90–105, 2004

[8] N. Tomasev, M. Radovanovic, D. Mladenic, and M. Ivanovic, “The role of hubness in clustering high-dimensional data,” IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 3, pp. 739–751, 2014.

[9] H. Kremer, P. Kranen, T. Jansen, T. Seidl, A. Bifet, G. Holmes, and B. Pfahringer, “An effective evaluation measure for clustering on evolving data streams,” in Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2011, pp. 868–876.

[10] P. Huang,Y. Huang,W.Wang, and L.Wang, ``Deep embedding network for clustering," in Proc. 22nd International. Conference. Pattern Recognition. (ICPR), Aug. 2014, pp. 1532-1537.

[11] P. Ji, T. Zhang, H. Li, M. Salzmann, and I. Reid, ``Deep subspace clustering networks," in Proc. Adv. Neural Inf. Process. Syst., 2017, pp. 23-32.

[12] W. Harchaoui, P. A. Mattei, and C. Bouveyron, ``Deep adversarial Gaussian mixture auto- encoder for clustering," in Proc. ICLR, 2017, pp. 1-5.

[13] F. Ye, C. Chen, and Z. Zheng, ``Deep autoencoder-like nonnegative matrix factorization for community detection," in Proc. 27th ACM Int. Conf. Inf. Knowl. Manag., Oct. 2018, pp. 1393- 1402.

[14] A. Rodriguez and A. Laio, ``Clustering by fast search and find of density peaks," Science, vol. 344, no. 6191, pp. 1492-1496, Jun. 2014.

[15] Z. Shu, X.Wu, H. Fan, P. Huang, D.Wu, C. Hu, and F. Ye, ``Parameter-less auto-weighted multiple graph regularized nonnegative matrix factorization for data representation," Knowl.- Based Syst., vol. 131, pp. 105-112,

[16] Sep. 2017 Y. Yang, H. Shen, Z. Ma, Z. Huang, and X. Zhou, “L21-norm regularized discriminative feature selection for unsupervised learning,” 07 2011, pp. 1589–1594

[17] J. Park, M. Lee, H. J. Chang, K. Lee, and J. Y. Choi, “Symmetric graph convolutional autoencoder for unsupervised graph representation learning,” in Proceedings of the IEEE International Conference on Computer Vision, 2019, pp. 6519–6528.

[18] S. Pan, R. Hu, G. Long, J. Jiang, L. Yao, and C. Zhang, “Adversarially regularized graph autoencoder for graph embedding.” in IJCAI, 2018, pp.2609–2615.

[19] C. Jie, Z. Jiyue, W. Junhui, W. Yusheng, S. Huiping, and L. Kaiyan, ``Review on the research of K-means clustering algorithm in big data," in Proc. IEEE 3rd Int. Conf. Electron. Commun. Eng. (ICECE), Dec. 2020 pp. 107-111.

[20] S. Ma, H. You, L. Tang, and P. He, ``An adaptive density peak clustering algorithm," J. Northeastern Univ. Natural Sci., vol. 43, no. 6, pp. 761-768, 2022