Predictive data Analytics using Multi-criteria Convolution Neural Architecture for effective Supply Chain Resilient Network of Micro Small Medium Enterprises

Ms. M. Sowmya Vani[1] , Dr. R. Mahammad Shafi[2]

[1]Research Scholar, Department of Computer Science, Bharathiar University, Coimbatore-641046, Tamilnadu, India.
[2]Research Supervisor, Department of Computer Science, Bharathiar University, Coimbatore-641046, Tamilnadu, India

Abstract

Predictive data analytics is a significant research focus in multiple Artificial Intelligence and Internet of thing incorporated supply chain management applications of the Micro Small Medium Enterprises. Performance of the predictive analytics depends on the quality of data clustering representation in terms of effectiveness and efficiency of supply chain data which contains high. Machine learning is a traditional approach of the predictive big data analytics to provide soft partition of supply chain data on forecasting the product demand to multiple customer base in the globalization market of MSME service industries but faces large complication due to increasing sparsity of data and increasing difficulties in distinguishing distance among data instances, especially it is highly complex in managing high dimensional data with complex latent distributions. In order to manage those complications, a new deep learning architecture has been illustrated in this paper. The proposed model is termed as Multi-criteria convolution Neural Network learning architecture for high dimensional data clustering and forecasting product demands. The current model explores deep feature towards for high cluster friendly representation. Initially Supply chain data is projected to missing value imputation towards imputation of the value to missing attribute as preprocessing process. Preprocessed supply chain data is employed to Linear discriminant analysis model for efficient feature extraction technique. Extracted features is employed to deep learning model for clustering on generation of objective function to produce maximum margin cluster. Those cluster further fine tuned to refine parameter of different layers of convolution neural network to ensure the minimum reconstruction error between feature max pooling layer and ReLu activation layer. Softmax layer minimizes the intra cluster compactness and inter cluster seperability in the feature space. Non parametric tuning has been enabled in the output layer to make the data instance in the cluster to be close to each other by determining the affinity of the data on new representation. It results significant increase in the clustering performance on the discriminative information's. Extensive experiments have been conducted on DatacoSupplychain dataset to compare current model with multiple conventional approaches. The experimental results show that current Multi-criteria convolution Neural Network learning architecture can achieve both effectiveness and good scalability on prediction outcomes and recommendation to high dimensional data supply chain network data .

Keywords: Micro Small Medium Enterprise, Deep Learning, Supply Chain Management, Predictive Data Analytics, High Dimensional Data, Data Clustering, Forecasting Product Demand

1. Introduction

Predictive Data Analytics (PDA) is becoming significant topic in the managing the supply chain data into effective data cluster which acquired on supply chain process of the manufacturing and service based industries. Supply chain data is exploited from planning, procurement, manufacturing, delivery and returns process by enabling to group an enormous amount of documents. However Clustering high dimensional data becomes sparse as the number of data instances required to represent any data distributions explores exponentially with the number of dimensions and distances among data instances leads to become complex to classify or cluster as dimensionality exploring high. The Complexity in handling high-dimensional data are omnipresent and abundant[1].

Machine learning model applies the phenomena of hubness [2] to extract the features from the attributes of the collected supply chain data for data clustering meanwhile it grows to curse of dimensionality challenges. Towards deep exploration of the features for predictive analytics with respect to product clustering and product forecasting on basis of customer product demand, deep learning based architectures has been explored to large extent to effective management of supply chain data. Deep learning architecture for predictive analytics of high dimensional supply chain data has been explored towards achieving soft clustering on achieving a new learning representation to transform supply chain entire process data into time series distributions[3].

Deep learning architecture is capable to alleviate the complication to produce effective and accurate the data clusters to high dimensional collected data. Especially optimizing of the convolution neural network architecture with feature learning criteria is becoming an active research focus to supply chain professional of Micro Small Medium Enterprises sectors in manufacturing and service based industries[4]. In this article, Multi-criteria Convolution Neural Network architecture has been proposed a deep learning approach for analyzing high dimensional supply chain data. Approach uses convolution neural network to map the extracted data into feasible feature space containing latent features on deep processing of convolution and max pooling layer.

Further fully connected layer of the model uses non -linear objective function represented as ReLu to eliminate the error. Finally soft max layer uses the classifier model to cluster the data features on basis of the product and customer location. Clustering helps to forecast the product demand on analysis of the product usage on refining parameters. Rest of the article is partitioned as follows, review of existing approaches for managing the supply chain data are described in section 2, the architecture of the current multi-criteria convolution neural network architecture is described in section 3 and experimental results and effectiveness of the current system is illustrated in section 4 using DatacoSupplychain dataset along performance comparison of proposed model against conventional approaches on multiple metric has been represented. Finally article is summarized with major findings in section 5.

2. Related works

In this section, high dimensional supply chain data composed of various process of the supply chain network is clustered on basis of location and price fluctuations using machine learning and deep learning approaches has been investigated in detail with respect to architectures for feature representations and similarity measures of the clustered data instances. Among those deep learning and machine learning techniques which produce good performance in terms of accuracy and effectiveness has been represented in detail and few techniques which performs equivalent to the current model is mentioned as follows

2.1. Artificial Neural Network For Predictive Data Analytics

In this architecture, Artificial Neural Network has been analyzed for managing the supply chain data on basis of clustering and forecasting. Artificial Neural Network belongs to machine learning provides the optimal trajectories on processing the high complex geographical data of the distribution and logistics data of the industries to the specified product or specified location or specified customer type. Initially kernel mechanism has been incorporated to extract the implicitly large features to organize the feature space[5]. Feature space is reconstructed as hierarchical random feature set. On employing partition rules, better feature representation into cluster has been achieved on retaining the close affinity among the data instance in the cluster.

2.2. Convolution Neural Network for Predictive Data Analytics

In this architecture, Convolution Neural Network has been employed for predictive data analytics of the supply chain data. Initially high dimensional supply chain data is transformed into low dimensional supply chain data space, further feature extraction mechanism is employed to extract product specific features on the data distribution of the supply chain network[6]. Extracted feature has been processed in the activation function to eliminate the reconstruction error to obtain the salient features for product and customer clustering. This model incorporate the various layer for sparse representation of the salient features and to eliminate the over fitting issue in the clusters by fine tuning of parameter with respect to certain criterion.

3. Proposed model

This section provides a design architecture of the current technique represented as multi criteria convolution neural network for predictive analytics of high dimensional supply chain data clustering on inclusion of parameter tuning of the deep learning layers to obtain the effective clusters for forecasting.

3.1. Data Pre-processing

Dataset preprocessing is carried out to transform the raw dataset into data mining specific format. In this work, preprocessing approaches contains the missing value imputation and dimensionality reduction process.

- **K Nearest Neighbour based Missing Value Imputation**

Missing Value Imputation is carried out using K Nearest Neighbour mechanism to high dimensional supply chain data. K Nearest Neighbour analyze the high dimensional data and computes missing attributes and value to it. It is carried out on computation of the centroids to the attribute in the specified data field through Euclidean distance function. Data instance of missing field is represented with K nearest value of the instance of the centroids [7].

- **Feature Extraction using Linear Discriminant Analysis**

Salient feature of the supply chain data is extracted using the linear discriminant Analysis (LDA) and it is capable of reducing the high dimensional supply chain data to lower dimensional supply chain data in the matrix form. Extraction model is employed to eliminate over fitting issue of the matrix [8]. Its objective is to obtain the feature space for clustering and forecasting. Feature space is represented as feature vector**.** The optimal feature vector towards clustering is obtained on employment of scatter matrix**.** The feature vector is provided in equation 1 as

Feature Vector for Attribute $f_i = \frac{1}{r} \sum_{i=0}^{n} D, (m - m_i)$ Eq.1

Total feature Vector for dataset $TF_i = \frac{1}{n} \sum_{x\varepsilon C}^{1} L, R$...........Eq.2

Scatter Matrix S_m for single feature vector $= \sum_{i=1}^{n} f_i \frac{mi}{r}$..........Eq.3

Scatter Matrix for multiple feature vector $S_{TM} = \frac{1}{2} \sum_{i=1}^{n} rF_j (f_i - TF)(f_i - TF_j)^T$..........Eq.4

where $(f_i - TF_j)$ is attribute variation

Projected Feature Matrix for data forecasting and clustering has been processed using scatter matrix [9]. Linear combination of the feature extracted is carried out using similarity computations[10].

3.2. Simulated Annealing for Feature Selection

Feature selection is a mechanism to estimate optimal features to generate effective data clusters. In this work, convolution neural network is implemented for forecasting and clustering of high dimensional supply chain data on extracted feature in this process. In this work, simulated annealing is employed to compute the new optimal feature space[11]. In order to obtain the optimal feature, population of the extracted features has to be initialized using Monte Carlo techniques on following constraints

Functional precedence constraints is mentioned as

$$F(\delta) = 1 - f/t \text{ ......Eq.5}$$

Functional precedence constraints produces the optimal feature space on managing the time and cost function based on neighbour selection of the features .Selection of the feature attribute for feature selection is carried out using acceptance function. Acceptance functions of each attribute to be optimal solution space is provided as .

$$F(T) = \exp(-fA/ct) \text{ ......Eq.6}$$

Attribute selected using simulated annealing model is presented using probability function towards effective feature representations[12] which can increase the discrimination of data instances on the similarity computation of the resultant feature space in data clustering task.

3.3. Multi-criteria Convolution Neural Network - Predictive Data Analytics

Multi-criteria Convolution Neural Network is deep learning architecture implemented for predictive data analytics. Predictive Data Analytics is carried out on the optimal feature space of the supply chain data. In this part, Convolution Neural Network employs hyper parameter tuning encomposing the activation function, objective function and loss function for cluster and forecasting the product demand. Initially convolution layer and max pooling is employed to maps the optimal features into latent features representation and these representations have been processed in fully connected layer to generate the forecasting and cluster results[13].

- **Max pooling layer**

Optimal feature space of the supply chain data is mapped to the cluster on basis of the forecasting and clustering constraints. Constraints formulated in the max pooling layer of the convolutional neural network and it is employed for mapping as follows

$$(C+ fP) = (C+ fp)^T R\ \alpha t\ (C+ fp) \text{ – First order Constraint}$$

$$(C+ fP) = C^T Rp + 2(fp)^T Pp + \alpha t (fq)^T P\ fp \text{ – Second order Constraint}$$

On incorporating the various levels of constraints fP which is orthogonal to p

$$(fp)^T R\ q\text{-}\lambda(fp)^T p \Rightarrow (fp)^T\ \beta t^{(} R\ q\text{-}\lambda\ q)=0........Eq.7$$

The feature representations which can increase the discrimination of the data instances during the similarity computation on data clustering is considered as outcome. The Hyper Parameter components of the Convolution Neural Network for clustering and forecasting the predictive data analytics.

Table 1: Hyper Parameter Component for the Convolution Neural Network

Hyper Parameter	Values
Optimal Layer Size	138
Model Learning Rate	0.01
Attribute Size	40
Number of Epoch in max layer	100
Activation Layer	ReLu
Maximum Sequence length of Instances	1000
Loss function	Cross entropy

- **Convolution Layer**

The Convolution layers of the model gathers the feature space through their inherent mechanism hierarchically from low level to more abstract features and learn the discriminative features with reduced parameters. It is represented in tensor form and it computes the weight of the each attribute using kernel function. Kernel Function determines the preference of the attribute on dot product approximation. Highly preferred attribute is considered as processing in fully connected layer

- **Batch Normalization**

Batch normalization is employed for faster convergence of the feature space with respect to epoch value to feature obtained using max pooling layer. Feature normalization acts the normalization of the activation function and its value. When BN is employed to the feature, it is not possible to influence feature weight propagation on the parameter scale. Thus, the learning rate of the model which manages the weights during decreasing of the attributes and it enables learning of the resultant features through ReLu function.

- **Activation Function**

Current architecture implements the rectified linear units (ReLU) activation function which introduces non-linearity to the features. Feature vector is processed on fully connected layer to various epoch layer with hyperparameter modified values to generate the appropriate cluster. Each activation function is to reduce or eliminate the over fitting and to enhance the system generalization by managing the output of the activation function with negative values with respect to decision function. It is implemented to trains the cluster to achieve high effectiveness than other functions.

- **Output Layer**

Output layer of the current multi-criteria convolution neural network contains the product clusters with customer information. Output layer contain the operation of the Soft max process and cross entropy mechanism to obtain the effective clusters. Effectiveness of the cluster is computed on basis of the determining the data affinity of the instance representation. Further Softmax layer reduces distance of the intra cluster and increase the distance of inter cluster.

Algorithm 1 : Multi Criteria Convolution Neural Network

Input: High Dimensional Supply Chain Data

Output: Product Cluster and Forecasting the product demand

Process

Data Pre- Process ()

Determine Missing value ()

Assign Centroids to Attributes containing missing values

Largest Nearest value of centroids to input to missing attributes

Feature Extraction

Feature _LDA ()

Compute Scatter matrix to attributes of the preprocessed supply chain data

Compute Covariance of Scatter matrix representing high discriminating features

Return feature Set F_s

Identify subspace for F_s

Compute Latent Feature on Subspace containing attributes

Feature Selection

Optimal features _Simulated Annealing

Mention precedence Constraints () to each attribute vector

Estimate feature fitness

Return Optimal Feature OF(s)

Multi criteria Convolution Neural Network ()

Max Pooling ()

Determine First order & Second Order Constraints to estimate weight features

Convolution Layer()

Estimate Sparse feature using kernel function

Fully Connected Layer ()

Batch Normalization()

ReLu_ Activation Layer()

Output Layer()

Softmax()--- Representation of the Product Cluster

Forecast _ cluster()

product demand forecasting

This function encourages the feature points on representative map to form cluster or become more discriminative in the particular cluster limit. Figure 1 represents the architecture of the proposed model.

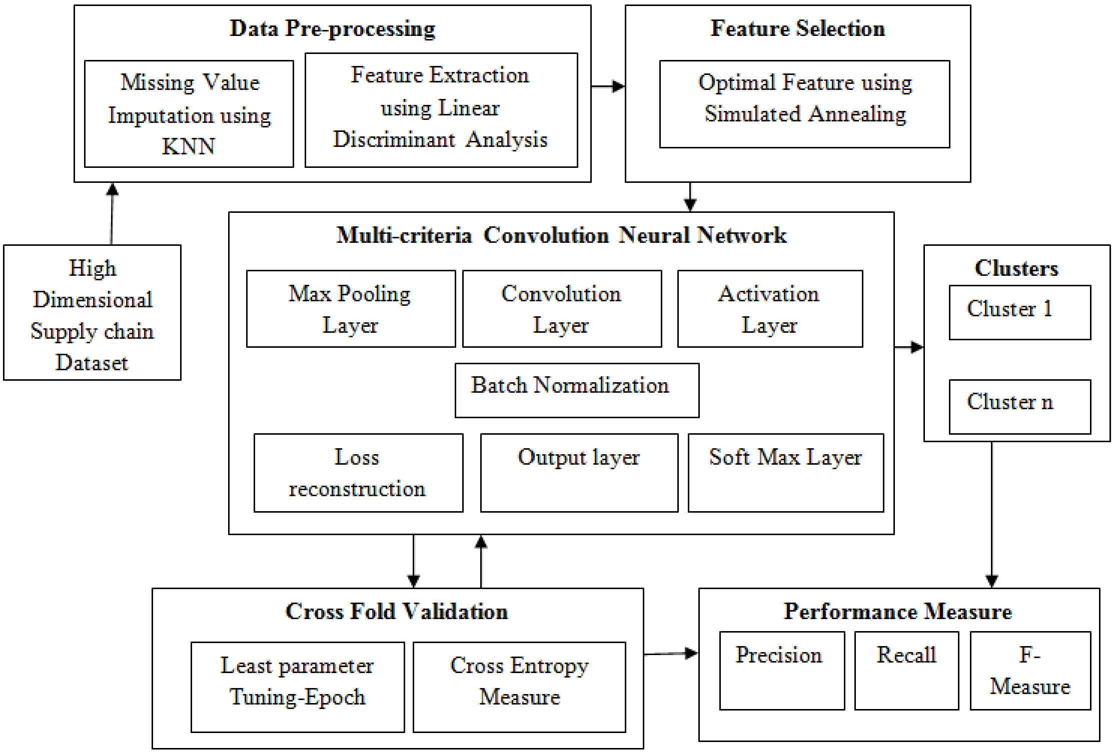

Figure 1: Architecture diagram

Training process of the Convolution Neural Network updates depending on the hyper parameter tuning of the traditional CNN architecture. Current model which guarantees that the learned representation provides the forecasting information and the data cluster of the collected data using learning representations which high data convergence rate . Further model is capable of learning the data instance with non linear dependency through the activation function and batch normalization process of model. Finally it reduces or eliminate the data sparsity challenges effectively on updating of the feature weight[14].

4. Experimental Results

Experimental analysis of proposed architecture using hyper parameter tuning on high dimensional data has been carried out on the various supply chain data of DatacoSupplychain dataset and Trademo dataset from supply chain management software of MSME industries. The performance of the current architecture is evaluated utilizing the precision, recall and F-measure to various real time dataset. The current architecture is experimented and evaluated using python technology. In this work, 60% of input dataset is used for model training and 20% of dataset is utilized for testing and remaining 20% of data to validate the current model through cross fold validation. Finally performance of the model is cross evaluated using 10 fold validation.

4.1. Dataset Description

Extensive experiments is carried out on DatacoSupplychain dataset and Trademo dataset from supply chain management software of MSME industries in order to compute the accuracy and efficiency of the model. Detailed description of the dataset is mentioned as follows

- **DatacoSupplychain dataset**

The data set contains data corpus of entire operation of the supply chain which describing the collection of each stage which is mostly frequently used bench mark dataset for many forecasting techniques[15].

- **Trademo.**

The data set contains geospatial descriptions of different types of supply chain information's. of the ecommerce applications[14].

4.2. Evaluation

The current architecture has been evaluated against the following performance measures against traditional machine learning and deep learning architectures. In this work, current architecture is evaluated using 10 fold validation to calculate the performance of forecasting on the clusters generated to specified dataset. The performance evaluation of the current deep learning architecture depends on the process of activation function on convolution layer, max pooling layer and fully connected layer representing the soft max function and cross entropy function.

- **Precision**

It is a computation of Positive predictive cluster value. It is further illustrated as the ratio of similar instances in the each supply chain cluster groups. In other way, Precision is termed as number of exact feature divided by the number of all returned feature space obtained. Figure 2 represents the performance assessment of the current architecture in terms of precision on DatacoSupplychain dataset and trademo dataset. Performance measures are suitable for computing the accuracy and efficiency of current architecture on forecasting to the generated cluster. Effectiveness of the model is obtained due to hyper parameter tuning

$$\text{Precision} = \frac{\text{True positive}}{True\ positive + False\ Positive} \text{....Eq.8}$$

Precision is computed in terms of true positive and true negative. True positive is a number of similar instances in the confusion matrix and false negative is number of real dissimilar instances in the confusion matrix [15]. Forecasting and clustering performance is characterized by high intra-cluster similarity and low inter-cluster similarity of the data instances in the clusters. It can be computed using recall measure.

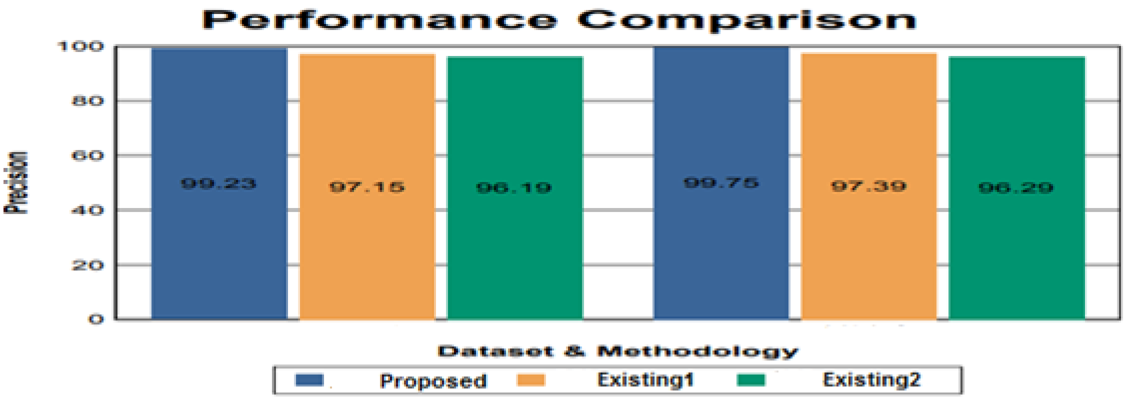

Figure 2: Performance analysis of Precision.

- **Recall**

Recall is the part of relevant feature instances points which extracted over the total amount of relevant instances of data cluster. The recall is the ratio of the relevant feature instances to the retrieved instances.

$$\text{Recall} = \frac{\text{True positive}}{True\ positive + False\ negative} \text{...Eq.9}$$

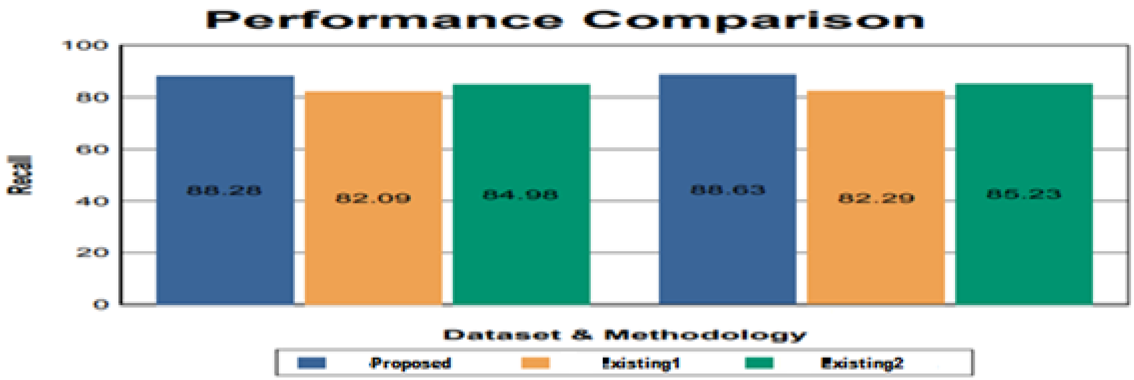

Figure 3: Performance analysis of Recall

True positive is a number of similar data points in the data and false negative is number of similar data points in the data. Figure 3 represents the performance assessment of the current architecture on recall measure along conventional approaches. Cluster quality depends on activation function of the architecture. F measure is a effective measure for computing the cluster quality.

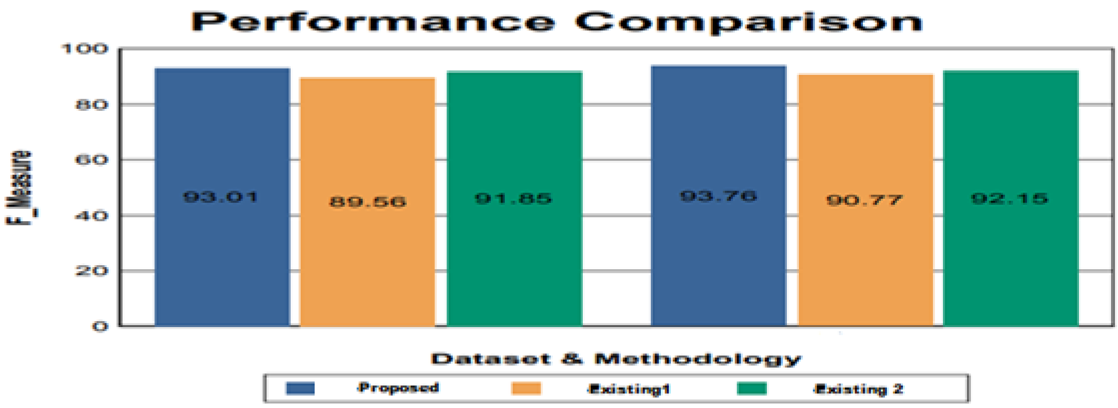

Figure 4: Performance analysis of the methodology on aspect of F- Measure

- **F measure**

It is the number of correct predictions to the supply chain data among total number of predictions to clusters.

Accuracy is given by

$$\frac{\text{True positive}+\text{True Negative}}{True\ positive+True\ Negative+false\ positive+False\ negative} \text{....Eq.10}$$

Although various data points may have multiple impact on cluster formation, they are likely to have the same impact on clustering. Figure 4 represents the performance of the current architecture in terms of f measure against conventional approaches for high dimensional supply chain data. Table 2 presents the performance value of the technique for cluster analysis.

Table 2: Performance Analysis of Deep learning architecture against Conventional techniques to DatacoSupplychain dataset

Technique	**Precision**	**Recall**	**F measure**
Multi-criteria Convolution Neural Network- Proposed	0.9918	0.8726	0.9781
Convolution Neural Network – Existing 1	0.9613	0.8418	0.9146
Artificial Neural Network - Existing 2	0.9518	0.8378	0.9085

On the other hand, the clustering method not only for detecting clusters in a given high dimensional data, but it is capable of detecting the underlying structure of the data distribution in general. It is naturally much more powerful, since they can handle nonhyperspherical clusters. Hyper-parameter is a very important component of the proposed deep fuzzy clustering models. In this grid search method with the extensive designed range to find the sensitive region of the hyper-parameters has been performed. In addition, the cross-validation has been used to twitter dataset alone to find the best value of the hyper-parameters

Table 3: Performance Analysis of Multi-criteria Convolution Neural Network against Conventional approaches to Trademo dataset

Technique	Precision	Recall	F measure
Multi-criteria Convolution Neural Network - Proposed	0.9978	0.8963	0.9376
Convolution Neural Network – Existing 1	0.9719	0.8419	0.8977
Artificial Neural Network Existing 2	0.9519	0.8413	0.8515

The evaluation of result is mentioned in the table 3 for Trademo dataset. It is observed that the current method is always better when compared to conventional approaches to supply chain resilience for supply chain management .

Conclusion

In this paper, Multi criteria Conventional Neural Network is designed and implemented as predictive data analytics for product demand forecasting and product clustering of the supply chain data to MSME industries. Initially data pre-processing is carried out for missing value imputation. Further Linear discriminant analysis is employed for feature extraction The reduced feature set has been processed using simulated annealing to generate optimal feature set for data clustering and forecasting. Current Architecture uses the Convolution Neural Network on hyper parameter tuning on the various layers to generate high representative cluster The architecture utilizes the optimal features in max pooling layer to generate sparse feature. Finally softmax layer and loss layer has been embedded in fully connected layer to produce the highly discriminative clusters and data forecasting. Cluster performance is evaluated using f measure as it proves the effectiveness on cluster. Finally predictive data analytics model proves that it is effective and high scalable to high dimensional supply chain data.

References

1. Belhadi, A., Kamble, S. S., Zkik, K., Cherrafi, A., & Touriki, F. E. (2020b). The integrated effect of Big Data Analytics, Lean Six Sigma and Green Manufacturing on the environmental performance of manufacturing companies: The case of North Africa. Journal of Cleaner Production, 252, 119903
2. Bottani, E., Centobelli, P., Gallo, M., Kaviani, M. A., Jain, V., & Murino, T. (2019). Modelling wholesale distribution operations: an artificial intelligence framework. Industrial Management & Data Systems, 119(4), 698–718.
3. 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.
4. 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.
5. Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 8, pp. 1798–1828, 2013
6. 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.
7. 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.
8. S.Praveen & R.Priya " A Deep Conceptual Incremental learning Based High Dimensional Data Clustering model- A Deep Learning Approach" in Turkish Journal of Computer and Mathematics, 2021
9. H. Liu, M. Shao, S. Li, and Y. Fu, "Infinite ensemble for image clustering," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, , pp. 1745–1754, 2016
10. 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
11. K. G. Dizaji, A. Herandi, C. Deng, W. Cai, and H. Huang, "Deep clustering via joint convolutional autoencoder embedding and relative entropy minimization," in 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, , pp. 5747–5756, 2017.
12. 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.

13. 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.
14. 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.
15. W. Harchaoui, P. A. Mattei, and C. Bouveyron, ``Deep adversarial Gaussian mixture auto-encoder for clustering," in Proc. ICLR, 2017, pp. 1-5.