IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-3, March 2019 E-ISSN: 2347-2693

Iterative Local Tangent Space Alignment with Adaptive Neighbourhood for Social Network Mining

S.K. Mishra[1*], A. Agarwal[2]

[1]Department of Computer Science, Accurate Institute of Management and Technology, Gr. Noida, India
[2]Department of Information Technology, KIET Group of Institutions, Ghaziabad, India

*Corresponding Author: sunilmishra.accurate @gmail.com, Tel.: +91-99995-73093

DOI: https://doi.org/10.26438/ijcse/v7i3.259265 | Available online at: www.ijcseonline.org

Accepted: 17/Mar/2019, Published: 31/Mar/2019

***Abstract*:** High dimensional data is a function of feature space inputs and noise. This makes it difficult to analyse and visualize the data, specially, in case of social network analysis where the aim is to analyse user's behaviour. Heterogeneous contents (text, audio, video, images etc.) make it more difficult to model. However, it is known that the actual feature vectors lie on much lower dimensions which can be obtained through non-linear manifold learning techniques. In this paper, we propose iterative local tangent space alignment with adaptive neighbourhood to extract the true low dimensional representation of data by sequentially aligning the tangent spaces and thereby reducing the overall reconstruction error. As the sub-manifold regions become linear, its neighbourhood size increases which leads to more information fusion. Extensive experiments on both synthetic and real world dataset proves that proposed method outperforms existing non-linear dimensionality reduction technique in both low dimensional representation and classification.

Keywords —Social Network Mining, Dimensionality reduction, Manifold learning, Local tangent space alignment, Adaptive neighbourhood

I. INTRODUCTION

The popularity of social network has increased exponentially and has become one of the most widely engaged platform across the globe [1][2][3][4]. Penetration to remote areas resulted in globally connected graph. Every minute, huge amount of heterogeneous content in posted online on such platforms. This makes it a very rich source of information where everybody can find something useful for them. With the same idea, multinational companies have made their way to social networks where they try and predict the behaviour of users. The major goals remain to identify whether a new product is going to be successful even before its actual production, the provided customer service is helpful or not, new way to marketing etc. However, this idea seems to be catchy and attractive, the shear amount of data and heterogeneous content restricts the existing state-of-the-art machine learning models to fully exploit this freely available information. As already known, most of the available data contains redundant information and hence, it can be thought of a manifold which embeds the actual low dimensional data having enough discriminative information for the underlying model.

Manifold learning assumes that the given high dimensional data do not contain unique or complement information in each dimension but there is redundancy and hence, actual data lie on a much lower dimensional space where data visualization, classification tasks can be done efficiently and accurately [5][6]. The process of extracting the low dimensional representation is known as dimensionality reduction. Traditional methods like principal component analysis (PCA), multidimensional scaling (MDS) etc. give true representation in case of linear manifold but as observed, not all manifolds follow linear properties especially in case of handwritten characters, spoken letters and social networks etc. To handle non-linear data, algorithms like isometric feature mapping (ISOMAP), local linear embedding (LLE), Laplacian eigenmap (LE), local tangent space alignment (LTSA) and many others have been proposed [7][8][9][10][11]. Among all, LTSA works on simple geometrical intuition and is easy to implement. Assume from a smooth Riemannian manifold $\mathcal{M} \in \mathbb{R}^D$, data samples $x_i \in X$ are drawn. The core idea of LTSA is to align the local tangent space around each x_i to determine the true connectivity between all manifold samples. In its core, LTSA assumes the following:

1. Given data is uniformly distributed.
2. Data in local neighbourhood of the manifold follow linear properties i.e. they lie in or close to a linear sub-space.

Due to above assumptions, the algorithms inherently also assume that the manifold's curvature is small, smooth and uniform which can be exploited in a single execution cycle but it fails to produce plausible inferences when (i) data follow non-uniform distribution; and (ii) sub-manifolds have large curvature.

To overcome these limitations, in this paper we propose iterative local tangent space alignment which solves the above problems by modelling LTSA with a feedback loop. Through multiple iterations on previously obtained results the algorithm efficiently aligns the large curvature sub-manifolds which were left partially aligned in previous execution. Instead of forcefully aligning the non-uniform samples, through iterations, the dissimilar observations are left in their previous aligned state which help the underlying model extract discriminative properties efficiently. Extensive experiments on synthetic and real world datasets confirm that proposed algorithm is capable of increasing the underlying machine learning model accuracy.

Rest of the paper is organized as follows: Section II contains a formal problem definition. In Section III our model of iterative LTSA with adaptive neighbourhood is explained. Section IV contains the experiments and results, and finally Section V concludes the paper.

1.1. Contribution

The proposed algorithm claims the following contributions:

1. Solves the problem of partially aligned tangent space through multiple iterations with feedback.
2. Adaptively increases the neighbourhood size with reduction in manifold's curvature.

II. PROBLEM DEFINITION

The social network can be modelled as a graph $G = \{V, E\}$, where each user or entity is represented by vertex $x_i \in V$, and connection between two vertices x_i, x_j is defined by a weighted edge between them like $e_{ij} = \delta(x_i, x_j) \in E$. In social network, generally each $x_i \in \mathbb{R}^D$ includes text, image, sound, videos etc. but most of these dimensions contain redundant or overlapping information that adversely affects the underlying machine learning model. This also leads to a complex solution as due to redundancy, data variance increases resulting in late or forced convergence. However, the true discriminative complement information lies on a much lower dimension $d \ll D$.

The given high dimensional social network data exhibit properties of a smooth Riemannian manifold where the low dimensional discriminative data is embedded in high dimensional space. Hence, the problem can be solved through any non-linear dimensionality reduction technique. However, due to heterogeneity in data, it follows non-uniform distribution leading to sub-manifolds' with large curvature which leads to incorrect predictions from existing state-of-the-art methods. The aim is to extract the true local geometrical data representation which contains highest discriminative information through iterative LTSA with adaptive neighbourhood to give accurate inferences.

III. ITERATIVE LTSA WITH ADAPTIVE NEIGHBOURHOOD

Given n data samples from a smooth Riemannian manifold $\{x_i\}_{i=1}^n \in \mathbb{R}^D$ that actually lies on a much lower dimension space $d \ll D$. It can be represented by

$$f: C \subset \mathbb{R}^d \to \mathbb{R}^D$$

where, C is a compact subset of $\mathbb{R}^d$ and f is the data generation function i.e.

$$x_i = f(\tau_i) + \eta_i$$

where, τ_i are original feature vectors or the lower dimensional complement information and η_i is redundant or noisy data. A linear manifold learning can be solved by minimizing the reconstruction error

$$min \parallel E \parallel = \underset{c,U,T}{\arg\min} \parallel X - (ce^T + UT) \parallel_F$$

where, $\parallel . \parallel_F$ defines the Frobenius norm on the matrix, $X = [x_1, x_2 \dots x_n]$, $T = [\tau_1, \tau_2 \dots \tau_n]$, $E = [\eta_1, \eta_2 \dots \eta_n]$, $e \in 1^n$ column vector and U contains orthonormal basis of intrinsic subspace. It can be solved by performing singular value decomposition (SVD)

$$SVD(X - \overline{X}e^T) = \mathbf{Q}\Sigma\mathbf{V}^T$$

where, $\mathbf{Q} \in \mathbb{R}^{D\times D}$, $\sigma \in \mathbb{R}^{D\times n}$, $\mathbf{V} \in \mathbb{R}^{n\times d}$ and $\overline{x}_i \in \overline{X} = \frac{1}{n}\sum_{j=1}^n x_j$ defines the centered matrix.

$$\Rightarrow UT = \mathbf{Q}_d \Sigma_d \mathbf{V}_d^T$$

The optimal U^* is given by eigenvectors $\mathbf{Q}_d$ corresponding to d largest singular values. Then, the linear manifold can be represented as

$$f(\tau) = \overline{X} + U * \tau$$

and the coordinate matrix Γ is given by

$$\Gamma = (U*)^T(X - \overline{X}e^T) = \mathrm{diag}(\sigma_1 \dots \sigma_d)\mathbf{V}_d^T$$

Γ contains the low dimensional representation of the given data. In case of non-linear dimensionality reduction, it is required to explore and exploit local linear region around each observation. The local linear structure can be extracted by representing each sample x_i with weighted linear sum of its neighbours N_i.

$$N_i = [x_{i,1}, x_{i,2} \dots x_{i,k}]$$

In its core LTSA works on the assumption that if the manifold is correctly unfolded then all tangent spaces will be aligned. The d dimensional sub-space for each N_i can be approximated by

$$\underset{x,\theta,Q}{\arg\min} \sum_{j=1}^{k} \parallel x_{ij} - (x + Q\theta_j) \parallel_2^2 = \underset{x,\Theta,\mathrm{Q}}{\arg\min} \parallel N_i - (xe^T + \mathrm{Q}\Theta) \parallel_2^2 \tag{1}$$

where, Q consists of d orthonormal columns and $\Theta = [\theta_1, \theta_2 \dots \theta_k]$. It holds the local tangent coordinates of neighbourhood data points. Further these local tangent coordinates will be aligned in lower dimensional space using affine transformations to obtain global coordinate system. Then,

$$\Theta_i = Q_i^T N_i (I - \frac{1}{k} ee^T) = [\theta_{i,1} \dots \theta_{i,k}]$$
$$\theta_{i,j} = Q_i^T (x_{i,j} - \overline{x}_i)$$
$$x_{i,j} = \overline{x}_i + Q_i \theta_{i,j} + \xi_{i,j}$$
$$\xi_{i,j} = (I - Q_i Q_i^T)(x_{i,j} - \overline{x}_i)$$

where, $\xi_{i,j}$ is the tangent reconstruction error. The global coordinate $\{\tau_i\}_{i=1}^{n}$ is constructed using the local coordinates $\theta_{i,j}$ where each $\tau_{i,j}$ should fulfil

$$\tau_{i,j} = \overline{\tau}_i + L_i \theta_{i,j} + \epsilon_{i,j}$$

for $i = 1 \dots n$ and $j = 1 \dots k$ defines each x_i's local neighbourhood.

$$\Rightarrow \Gamma_i = \frac{1}{k} \Gamma_i ee^T + L_i \Theta_i + E_i \qquad (2)$$

where, $\Gamma_i = [\tau_{i,1} \dots \tau_{i,k}]$ and $E_i = [\epsilon_{i,1} \dots \epsilon_{i,k}]$ is the local reconstruction error.

$$E_i = \Gamma_i (I - \frac{1}{k} ee^T - L_i \Theta_i) \qquad (3)$$

The optimal L_i for a fixed Γ_i is given by

$$L_i = \Gamma_i (I - \frac{1}{k} ee^T) \Theta_i^+ = \Gamma_i \Theta_i^+$$

$$\therefore E_i = \Gamma_i w_i \text{ where, } w_i = (I - \frac{1}{k} ee^T)(I - \Theta_i^+ \Theta_i)$$

where, Θ_i^+ represents pseudo inverse of Θ_i.

$$\Rightarrow \sum_{i=1}^{n} \| E_i \|_F^2 = \| TSW \|_F^2$$

where, $S = [s_1 \dots s_n]$ is selection matrix such that $\Gamma s_i = \tau_i$ and $W = \mathrm{diag}(w_1 \dots w_n)$. Constraint $\Gamma\Gamma^T = I_d$ helps finding unique Γ. Then, the vector e becomes the eigenvector of $B = SWW^T S^T$ with respect to eigenvalue zero hence, optimal Γ is given by the eigenvectors corresponding to 2nd to $d + 1$ smallest eigenvalues of B.

While, single LTSA execution aligns the tangent spaces when the samples are uniformly distributed and the manifold's curvature is small but it give partial alignment when either of these criteria fails. However, executing LTSA in multiple iterations with feedback loop and adaptive neighbourhood effectively solves the problem. In the beginning, the iterative LTSA begins with small neighbourhood assuming large curvature resulting in large reconstruction error $\epsilon_{i,j}$. As the tangent spaces are aligned, the curvature of low dimensional representation is reduced and hence, it can now accommodate more neighbours than the previous cycle which also reduces $\epsilon_{i,j}$. The process is repeated until the reduction in $\epsilon_{i,j}$ is no longer significant as defined by threshold.

Algorithm 1: Iterative LTSA with adaptive neighbourhood

Input: $\boldsymbol{X} \in \mathbb{R}^D$ **:** Data
***nn*:** Initial Neighbours
***nnStep*:** Neighbourhood Step
Δ: Error Threashold
Output: $\boldsymbol{X'} \in \mathbb{R}^d$ **:** Low dimensional representation

pErr $\leftarrow \infty$
while *true* **do**
  for *i* $\leftarrow$ 1 to *n* **do**
    $N_i \leftarrow$ find_knn(x_i, nn);
    $\overline{x}_i \leftarrow \frac{1}{nn} \sum_{j=1}^{nn} x_{N_{i,j}}$
    $t \leftarrow (N_i - \bar{x}_i e^T)^T (N_i - \bar{x}_i e^T)$
    $[g_1 \dots g_d] \leftarrow$ pca(t,d);
    $G_i \leftarrow [e/\sqrt{k}, g_1 \dots g_d]$;
    $B(I_i,I_i) \leftarrow B(I_i,I_i) + I - G_i G_i^T$;
  end
  /* Get global coordinates
  $\Gamma \leftarrow$ svd(*B,d,'small'*);
  /* Mean of reconstruction error
  nErr $\leftarrow$ mean(eqn.3);
  /* Break when improvement is insignificant
  if *(pErr – nErr)* < Δ **then**
    break;
  end
  nn $\leftarrow$ *nn* + *nnStep*; // Increase *nn*
  X $\leftarrow \Gamma$; // Feedback loop
  pErr $\leftarrow$ *nErr*; // Replace previous error
end
X' $\leftarrow$ *X*;
return *X'*;

IV. EXPERIMENTS AND RESULT

Extensive experiments for non-linear dimensionality reduction have been performed on 1 synthetic and 3 real dataset by employing LTSA, ISOMAP, LLE, Laplacian eigenmap and proposed iterative LTSA with adaptive neighbourhood. Two classifier models kNN and SVM were trained with each method and were compared among themselves [12].

IV.I Elevated Swiss roll

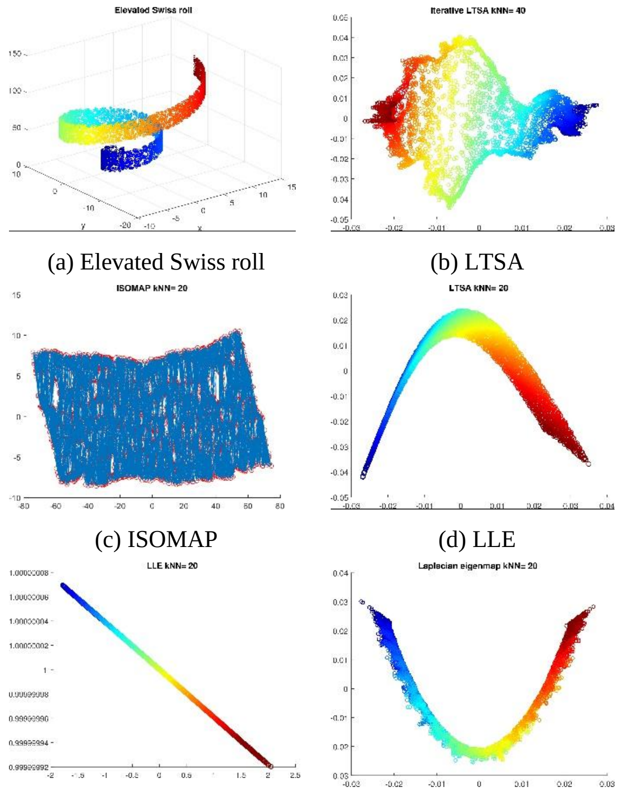
Elevated Swiss roll is a synthetic data set containing 4000 sample points belonging to $\mathbb{R}^3$ but as we can see, the data points actually lie on a flat $\mathbb{R}^2$ strip. The used data differs from classical Swiss roll by having a variation in z axis as shown in fig. 1(a). Here, we have compared state-of-the-art non-linear dimensionality reduction techniques. As evident from the result, iterative LTSA with adaptive neighbourhood is able to capture most accurate intrinsic geometry as compared to others. The proposed algorithm started with $kNN = 5$; went up to 40 by step of 5 neighbours in each iteration.

To make a fair comparison, every dimensionality reduction was configured to work with kNN = 20. The basic LTSA was able to capture the most outer data samples colour coded red and rest of them were merged to a single blob. Specially, in the case of inner most data points as shown in fig. 1(b). ISOMAP was able to flatten the structure but due to variation in all ambient dimensions, it resulted into an overlapping strip as shown in fig. 1(c). ISOMAP proved competent enough to find the true intrinsic geometrical structure but failed to produce its correct lower dimensional representation. Fig. 1(d) shows the output of LLE, in case of clustering, this result proves viable but overlapping structure looses the information hidden in other data points. Laplacian eigenmap as shown in fig. 1(e) produced a sub-optimal curved strip with data points neatly arranged in respective clusters. However, even Laplacian proved inefficient to obtain its true representation. In the end, fig. 1(f) shows result of proposed algorithm. As evident, iteration with adaptive neighbourhood is able to perform far better than other algorithms. It provides a minimal overlapping representation which identifies the true intrinsic geometry of data.

IV.II Facebook metrics

The data is captured from Facebook in 2014 for a leading cosmetic brand to predict the impact of online posts [1]. It includes 500 observation spanned over 18 attributes among which only 7 features constitute input space while rest 11 are prediction values. The data had 4 missing values which were replaced by mean of their respective dimension. Input features holds information like data/time of post, unique post identification, content of the post, post type/category, and paid post. Based on these values, the corresponding performance predictions for lifetime post total reach, impression, user engagement, post consumption, likes, comments, share etc. needs to be made. As already known, not all input dimensions constitute in same proportion for final inference, we performed existing and proposed non-linear dimensionality reduction and compared their performance based on minimalistic kNN classifier, as well as complex SVM classifier.

As evident from fig. 2 and 3, iterative LTSA with adaptive neighbourhood based classifiers outperform every other model. As the proposed method iteratively aligned the local tangent spaces which adaptively increased the linear neighbourhood to further enhance the notion of similarity and dissimilarity between two entities.



(a) Elevated Swiss roll (b) LTSA

(c) ISOMAP (d) LLE

(e) Laplacian eigenmap (f) Iterative LTSA

Figure 1: Dimensionality Reduction

Table 1: Mean Error (Facebook Metrics)

Algorithm	kNN	SVM
Iterative LTSA	23.1098	21.4131
Full Features	29.6827	28.1373
ISOMAP	28.2794	26.9158
LTSA	28.1546	26.5894
LLE	27.8523	26.2160
Laplacian	27.3032	25.9396

As compared to kNN, SVM produced more accurate results. Table 1 shows the mean error percentage of all 11 performance attribute. The proposed method proved to increase the model's accuracy by ≥4%. Model created on all 7 input dimensions failed to give accurate results, while ISOMAP and LTSA gave comparable results. Similarly, LLE and Laplacian being similar model, produced approximately same result.

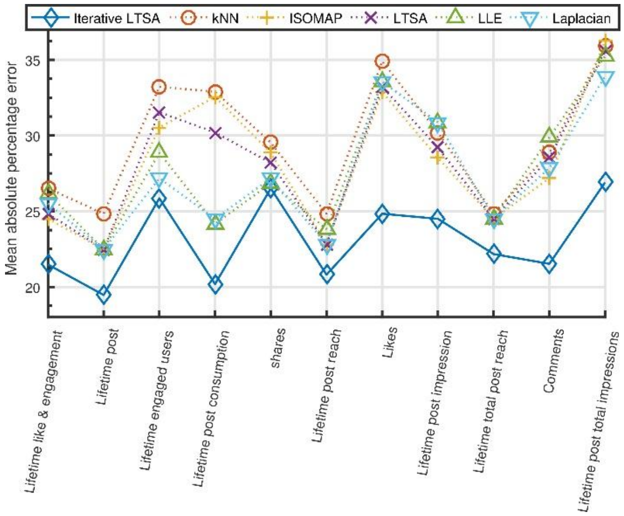


Figure 2: Facebook kNN classification

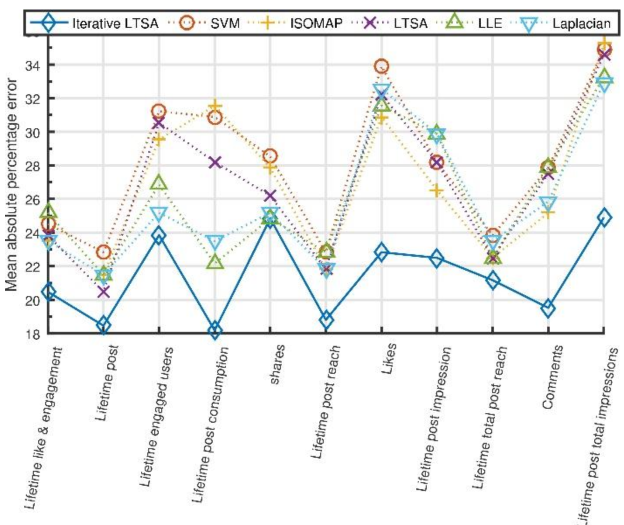


Figure 3: Facebook SVM classification

IV.III Natural images

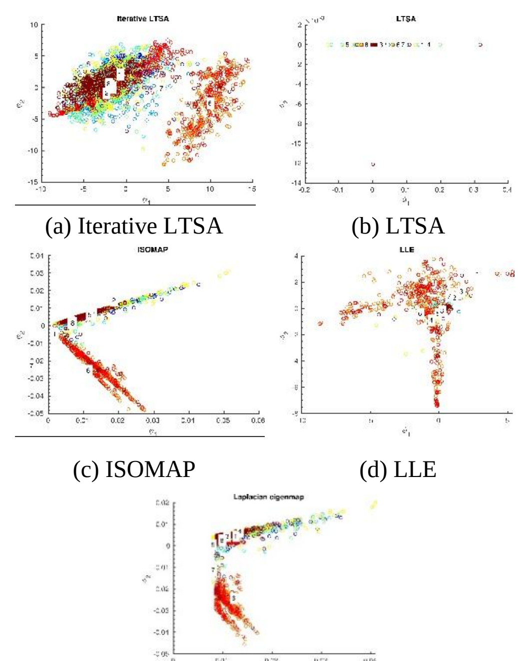
Natural image dataset contains 7599 images of 8 different categories (airplane, car, cat, dog, flower, fruit, motorbike and person) [13]. Snapshot of the data is shown in fig. 4. The dataset was divided into two groups train and test containing 3800 and 3799 instances respectively. Both train and test group contain uniform observation from all categories. As the images were not of uniform resolution, in pre-processing, every image was resized to 28x28x3. The existing state-of-the-art non-linear dimensionality reduction techniques, LTSA, ISOMAP, LLE and Laplacian eigenmap are employed along with proposed method to reduce the data dimensions to 2 as shown in fig. 5. As evident, the proposed iterative LTSA with adaptive neighbourhood gives best low dimensional representation where every category can be easily visualized. Next in performance were ISOMAP and Laplacian eigenmap. Basic LTSA and LLE failed to extract the true intrinsic geometrical properties and hence an overlapping blob is given as their low dimension representation.

Both kNN and SVM classifiers were trained on all five methods using their low dimensional representation of training set. Table 2 shows the mean error of all methods for test set prediction. SVM being a more sophisticated classifier gave accurate results than kNN. Based on accuracy parameter, the proposed algorithm outperformed other four methods by a comfortable margin.

Table 2: Mean Error (Natural Images)

Algorithm	kNN	SVM
Iterative LTSA	12.6389	11.4668
Full Features	24.1748	22.5000
ISOMAP	13.1548	13.6379
LTSA	25.3917	20.6610
LLE	22.3851	16.5000
Laplacian	13.3259	12.6412

Figure 4: Natural images (8 categories)



(a) Iterative LTSA (b) LTSA

(c) ISOMAP (d) LLE

(e) Laplacian eigenmap

Figure 5: Natural images

IV.IV Fashion MNIST

Fashion MNIST dataset is a collection of 60000 training and 10000 test images from 10 different fashion categories (top, Trouser, Pullover, Dress, Coat, Sandal, Shirt, Sneaker, Bag, Ankle boot) [14]. Each image consist of 28 x 28 grayscale pixels. Fig. 6 shows a snapshot of images included in this dataset. Among all five non-linear dimensionality reduction technique, the proposed iterative LTSA with adaptive neighbourhood gave true low dimensional representation as shown in fig. 7a. The different clusters are easily identifiable. Basic LTSA gave blob like low dimension representation which remained least accurate as shown in fig. 7b. The other three methods ISOMAP, LLE and Laplacian eigenmap gave good results but were not as accurate as proposed method.

Table 3: Mean Error (Fashion MNIST)

Algorithm	kNN	SVM
Iterative LTSA	6.1340	5.9020
Full Features	10.7240	10.0040
ISOMAP	7.1940	6.8340
LTSA	9.8460	9.2380
LLE	8.5760	8.0340
Laplacian	7.4510	6.9760

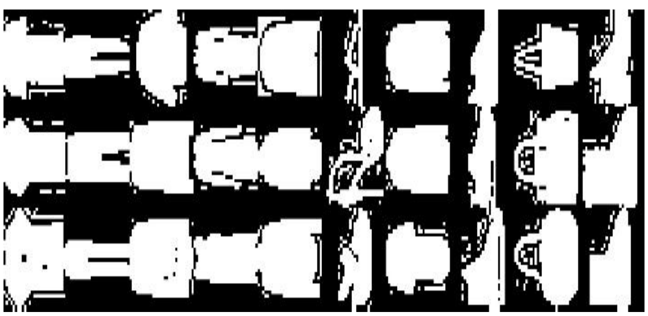
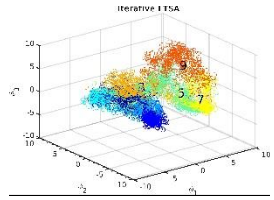
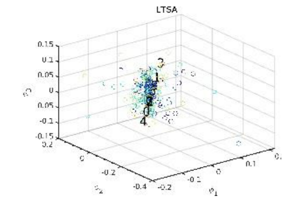


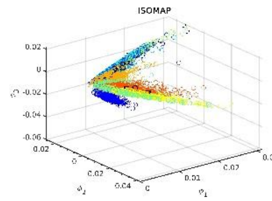
Figure 6: Fashion MNIST (10 categories)



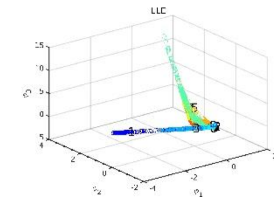
(a) Iterative LTSA



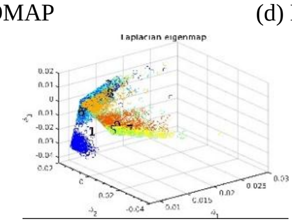
(b) LTSA



(c) ISOMAP



(d) LLE



(e) Laplacian eigenmap

Figure 7: Fashion MNIST

kNN and SVM classifier models were trained with 60000 observations as obtained from all five non-linear dimensionality reduction methods and original given data. Overall the accuracy of SVM remained more than kNN classifier. Among all classifiers, model trained with proposed method gave > 93% accurate classification results which is highest among all methods used. Full dimension model gave ≈ 90% accuracy which was least among all followed by Laplacian, ISOMAP, LLE and LTSA.

Remarks: Thus, the above experiments show that proposed method is able to project given higher dimensional non-uniformly distributed heterogeneous data to its true low dimensional feature space with minimal information loss. In the process of learning impact of social network, the proposed method proved to be much accurate than other existing methods. Iterative LTSA with adaptive neighbourhood remains capable of handling large manifold curvatures by partially aligning them until they get completely aligned which ultimately transforms the data to a linear space. This makes visualization, modelling, classification, regression etc. on the data easy and accurate.

V. CONCLUSION

In social networks, the assumption that given data is homogeneous and uniformly distributed, leads to failed or sub-optimal low dimensional representation. Large manifold curvature restricts the local linear region thus reducing number of neighbours which increases the local reconstruction error. Iterative LTSA reduces the curvature partially in each execution by aligning the tangent spaces finally leading to a global aligned tangent space. Adaptive neighbourhood exploits the property of small curvature by increasing the number of neighbours in each iteration thus reducing the local reconstruction error. Experiments on both synthetic and social network data show that proposed method extracts the true low dimensional representation that ultimately increases the accuracy of underlying machine learning model. Experimental evidences supports the fact that proposed method outperforms the existing state-of-the-art nonlinear dimensionality reduction method by huge margin. Algorithms being generic in nature, can be adapted to various other scenarios.

REFERENCES

[1] S. Moro, P. Rita, and B. Vala, "*Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach*," Journal of Business Research, vol. 69, no. 9, pp. 3341–3351, 2016.

[2] J. Surma and A. Furmanek, *"Data mining in on-line social network for marketing response analysis,"* in 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, Oct 2011, pp. 537–540.

[3] L. del Carmen Contreras Chinchilla and K. A. R. Ferreira, *"Analysis of the behavior of customers in the social networks using data mining techniques,"* in 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Aug 2016, pp. 623–625.

[4] Hussain, Sadiq, et al. *"Big data and learning analytics model."* International Journal of Computer Sciences and Engineering 6.7 (2018): 654-663.

[5] T. Jia, M. Jian, L. Wu, and Y. He, *"Modular manifold ranking for image recommendation,"* in 2018 IEEE Fourth International Conference on Multimedia Big Data (BigMM), Sep. 2018, pp. 1–5.

[6] W. Zhao, S. Tan, Z. Guan, B. Zhang, M. Gong, Z. Cao, and Q. Wang, *"Learning to map social network users by unified manifold alignment on hypergraph,"* IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 12, pp. 5834–5846, Dec 2018.

[7] J. B. Tenenbaum, V. De Silva, and J. C. Langford, *"A global geometric framework for nonlinear dimensionality reduction,"* science, vol. 290, no. 5500, pp. 2319–2323, 2000

[8] L. K. Saul and S. T. Roweis, *"An introduction to locally linear embedding."*

[9] M. Belkin and P. Niyogi, *"Laplacian eigenmaps for dimensionality reduction and data representation,"* Neural computation, vol. 15, no. 6, pp. 1373–1396, 2003.

[10] Z. Zhang and H. Zha, *"Principal manifolds and nonlinear dimensionality reduction via tangent space alignment,"* SIAM JOURNAL ON SCIENTIFIC COMPUTING, pp. 313–338, 2004.

[11] Vidya, G., and S. Omprakash. *"Survey on Recent Researches on High Level Image Retrieval."* International Journal of Computer Sciences and Engineering (2016): 72.

[12] C. Cortes and V. Vapnik, *"Support-vector networks,"* Machine learning, vol. 20, no. 3, pp. 273–297, 1995.

[13] P. Roy, S. Ghosh, S. Bhattacharya, and U. Pal, *"Effects of degradations on deep neural network architectures,"* arXiv preprint arXiv:1807.10108, 2018.

[14] H. Xiao, K. Rasul, and R. Vollgraf. (2017) Fashion-mnist: a novel image dataset for benchmarking machine learning algorithm.

Authors' Profile

S K Mishra pursed his Bachelor of Technology in Computer Science from Dr. A.P.J. Abdul Kalam Technical University Uttar Pradesh, Lucknow in year 2005 and Master of Technology in year 2012. He is currently pursuing Ph.D. from IFTM University, Moradabad under the supervision of Professor Ajay Agrawal and working as an Assistant Professor in the Department of Computer Sciences, Accurate Institute of Management and Technology, Greater Noida since 2012. Previously he worked as a lecturer in KNIT, Sultanpur from 2005 to 2007 and as an Assistant professor in UCER, Greater Noida from 2007 to 2012, making it approximately 14 years of academic experience. He has published many research papers in reputed journals and conferences. His main research work focuses on social network mining, web mining and literature mining etc.

Ajay Agarwal received his B. Tech. degree from IET Lucknow (a constituent college of AKTU Lucknow), in year 1979. He completed his Master of Engineering (Computer Science and Engineering) with honors from M.N.N.I.T. Allahabad in year 1995 and Ph.D. from I.I.T. Delhi in year 2006. He is a versatile professional with a strong academic background. His area of research includes mobile adhoc networking, soft computing, data base management system. He has received many awards. He has published many papers in reputed journals and conferences. He is an author of a book. Presently he is working as Professor in Information Technology Department at KIET Group of Institutions, Ghaziabad.

