International Journal of Research in Advent Technology, Vol.7, No.4S, April 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

Dimensionality Reduction Techniques- Survey

Dr. C. Rajakumar
Associate Professor,
Department of ECE, Anurag Group of Institutions, Hyderabad

Abstract: In real world scenario, storing data, audio, image and video is big challenging problem. By reducing the dimension of the data from the original dimension to lower dimension leads to good visualization, less computation time and faster execution time. Lot of dimensionality reduction techniques exists and classified into two categories' namely feature selection and feature extraction. Feature selection is removal of irrelevant and redundant data thereby reducing in computation time and increasing accuracy. A data set may contain many futures and all the features are not important and only the important feature is considered and features that are not important is discarded. Feature extraction or projection is mapping higher dimensionality data into lower dimensional data. The mapping may be linear or nonlinear. In this paper comparison of all dimensionality reduction is employed and best dimensionality reduction techniques is auto encoder-based approach.

Keywords: Principal Component Analysis, Linear Discriminant Analysis, Autoencoder

1. INTRODUCTION:

Large scale data with higher dimension is a big challenging problem in machine learning. Learning algorithm and making model is a big challenging task with or without the presence of noise. Feature selection is selecting important features and discarding other irrelevant features. The classification of feature selection is unsupervised [1,2,3,4,5] , supervised[6,7,8,9] and unsupervised[10,11,12] . Dimensionality reduction techniques are reducing the dimension of original data to lower dimension. The techniques may be linear or nonlinear. It is achieved by subset of existing features. Reduced computation time, data compression and reduce storage space are the advantages of dimensionality reduction techniques. Loss of information is unavoidable in reducing dimension. Principal component analysis, Linear discriminant analysis, Non-negative factorization and auto encoder are some of the dimensionality reduction techniques.

2. PRINCIPAL COMPONENT ANALYSIS:

This technique extracts features in terms of principal components. The first principal component corresponds to maximum variance in the data and the second principal component corresponds to variance in a direction that is perpendicular to first principal component analysis. PCA is a statistical based approach. It transforms a set of correlated variables into linearly uncorrelated variables. Assume that O observations in the data and variables are p, number of principal components is min(O-1, p).PCA is linear orthogonal transformation. The steps involved in principal component analysis is 1) Formation of mean centered data 2) Normalize the data 3) Eigenvector and eigenvalue calculation 4) Formation of principal components.

3. NON-NEGATIVE MATRIX FACTORIZATION

It factors the non-negative into products of two non-negative matrices. Let the matrix be A.A=WH where W and H are non-negative matrices. WH is lower rank approximation to A. Initially random values are given for W and H. Based on interactive method, W and H values are generated. In some cases, the algorithm converges to lower rank than k.

This is a rectangular matrix. The product of W and H is equal to A. Non-negative matrix factorization (Cichochi & Anh 2009) is originally known as non-negative rank factorization or matrix factorization. NMF has been analyzed for more than 25 years. NMF can be classified into two types. They are optimization based methods (Chu & Lin 2008) and geometry based methods (Yuksel et al 2011). The restriction on k is that NMF is generally not unique. The factor matrices of NMF are always non negative. NMF is a non-supervised technique that can be used for future learning. To quantify the approximation error, normally cost functions are used. A simple update rules for W and H by minimizing the divergence between A and WH. W and H matrices may be sparse. This means columns of W and or rows of H may be zero. NMF is the one of the important methods to approximate the measured data. MATLAB software uses alternate least square algorithm or multiplicative update algorithm to factorize NMF. In document analysis, each document is stored as vector. Each element of vector indicates count of a term appearing that document. In image processing, each vector in a

matrix represents an image. A matrix represents a collection of images. Each element of the vector represents color of pixel. NMF extracts facial parts from face images.

In consumer market, history of each customer or rate of each item can be represented by non-negative matrix. In gene analysis, various experiments are carried out to identify a particular gene. All the experimental results are represented by a non-negative matrix. When the measurement is made on the height and weight of certain people, it can be represented by non-negative matrix. An example is given below that indicates the height and weight of four persons. The main drawback of NMF algorithm is initialization of W and H. Some algorithms initialize both W and H. Some algorithms in NMF that initialize only W. Some of the application of NMF is Face and object recognition (Soukup & Bajla 2008), Watermarking in images (Lee et al 2009) Classification of texture (Sandler & Lindenbaum, 2011), Image denoising (Mairal et al, 2010)

4. LINEAR DISCRIMINANT ANALYSIS

2D-FLDA is directly applicable to the given matrix whereas one Dimensional Fisher Linear Discriminant Analysis(1DFLDA) operates on vectors. Figure 3.1 shows the block diagram of 2D-FLDA for two classes. It separates the classes by finding mean of each class, within class scatter matrix, between class scatter matrix, projection vectors and optimal projection vectors. Optimal projection vectors retain only first few vectors from projection vectors. Remaining less significant vectors from projection vectors are eliminated.

5. AUTOENCODER :

To make a representation for a set of data in a unsupervised manner, autoencoder is used. It tries to reconstruct the input from the learned representation. It consists of encoder and hidden layer and decoder. First layer is input layer and middle layer is hidden layer and third layer is output layer. Input layer and output layer has same number of nodes. Activation function used in autoencoder is sigmoid and rectified linear unit. Reconstruction error between input and output should be minimized. Vector space spanned by p-principal components and vector space spanned by weights of hidden layer with p nodes are same. Autoencoder uses backpropagation algorithm like stochastic gradient descent, conjugate gradient descent, steepest gradient descent. Regularizer term is added to avoid overfitting.

6. CONCLUSION:

All dimensional reductional techniques convert higher dimensional data to lower dimensions. Auto encoder posses high computation time. But in terms of reconstruction it exactly reproduces the original data compared with all other methods and accuracy is very high.

REFERENCES

[1] L. Wolf, A. Shashua, Feature selection for unsupervised and supervised inference: The emergence of sparsity in a weight-based approach, The Journal of Machine Learning Research 6 (2005) 1855–1887.

[2] Z. Zhao, H. Liu, Spectral feature selection for supervised and unsupervised learning, in: Proceedings of the 24th international conference on Machine learning, ACM, 2007, pp. 1151–1157.

[3] F. Nie, H. Huang, X. Cai, C. H. Ding, Efficient and robust feature selection via joint ?2, 1-norms minimization, in: Advances in neural information processing systems, 2010, pp. 1813–1821.

[4] J. Li, Z. Chen, L. Wei, W. Xu, G. Kou, Feature selection via least squares support feature machine, International Journal of Information Technology & Decision Making 6 (04) (2007) 671–686.

[5] Z. Zhao, H. Liu, Semi-supervised feature selection via spectral analysis., in: SDM, SIAM, 2007, pp. 641–646.

[6] Z. Xu, I. King, M. R.-T. Lyu, R. Jin, Discriminative semi-supervised feature selection via manifold regularization, Neural Networks, IEEE Transactions on 21

(7) (2010) 1033–1047. [7] P. Wang, Y. Li, B. Chen, X. Hu, J. Yan, Y. Xia, J. Yang, Proportional hybrid mechanism for population based feature selection algorithm, International Journal of Information Technology & Decision Making (2013) 1–30.

[8] D. Cai, C. Zhang, X. He, Unsupervised feature selection for multi-cluster data, in: Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2010, pp. 333–342.

[9] J. G. Dy, C. E. Brodley, Feature selection for unsupervised learning, The Journal of Machine Learning Research 5 (2004) 845–889.

[10] Y. Yang, H. T. Shen, Z. Ma, Z. Huang, X. Zhou, l2, 1-norm regularized discriminative feature selection for unsupervised learning, in: IJCAI

Proceedings-International Joint Conference on Artificial Intelligence, Vol. 22, Citeseer, 2011, p. 1589.

[11] E. R. Hruschka, E. R. Hruschka Jr, T. F. Covoes, N. F. Ebecken, Bayesian feature selection for clustering problems, Journal of Information & ˜ Knowledge Management 5 (04) (2006) 315–327.

[12] R. Liu, R. Rallo, Y. Cohen, Unsupervised feature selection using incremental least squares, International Journal of Information Technology & Decision Making 10 (06) (2011) 967–987.

[13] Cichocki, A & Anh-Huy, P 2009, 'Fast local algorithms for large scale nonnegative matrix and tensor factorizations', IEICE transactions on fundamentals of electronics, communications and computer sciences, vol. 92, no. 3, pp. 708-721.

[14] Yuksel, A, Akarun, L & Sankur, B 2011, 'Hand vein biometry based on geometry and appearance methods', IET computer vision, vol. 5, no. 6, pp. 398-406.

[15] Soukup, D & Bajla, I 2008, 'Robust object recognition under partial occlusions using NMF', Computational Intelligence and Neuroscience.

[16] Lee, H, Cichocki, A & Choi, S 2009, 'Kernel nonnegative matrix factorization for spectral EEG feature extraction', Neurocomputing, vol. 72, no. 13, pp. 3182-3190.

[17] Sandler, R & Lindenbaum, M 2011, 'Nonnegative matrix factorization with earth mover's distance metric for image analysis', IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 8, pp. 1590-1602.

[18] Mairal, J, Bach, F, Ponce, J & Sapiro, G 2010, 'Online learning for matrix factorization and sparse coding', The Journal of Machine Learning Research, vol. 11, pp. 19-60.