Performance Analysis of Dimensionality Reduction using PCA, KPCA and LLE for ECG Signals

J Haris Mita¹, C Ganesh Babu² and M Gowri Shankar³

1 Student, Bannari Amman Institute of Technology, Tamilnadu, India
2 Professor, Bannari Amman Institute of Technology, Tamilnadu, India
3 Assistant Professor, Gnanamani College of Technology, Tamilnadu, India

harismita.es19@bitsathy.ac.in

Abstract. Machine Learning is a predominantly developing field in the current scenario due to spontaneous growth in size of data. Dimensionality Reduction is significant step used in machine learning. The Imprecation of dimensionality refers to the difficulty that comes up when working with high dimensional data. The Proposed work developed to reduce the risk of high dimensional data representation in the form of low dimensional data representation. This paper presents the dimensionality reduction practices such as Principal Component Analysis (PCA), Kernel PCA and Locally Linear Embedded (LLE).

Keywords: Machine Learning, Dimensionality Reduction, PCA, LLE, Kernel PCA.

INTRODUCTION
Dimensionality Reduction is the modification of large-scale input into reasonable low-dimensional output representation. The reduced data is preferably parallel to its underlying dimensionality. Dimensionality Reduction is one of the dominant techniques in several areas since it enable high dimensional data to be visualized, sorted and compressed by weakening the dimensional curse.

Dimensionality Reduction is based on two methods such as linear and non-linear. A significant number of non-linear methods have currently been used to reduce dimensionality. All these techniques are based on the fact that data is contained in a complex low dimension manifold embedded into the high dimension space. The goal is to identify and to remove the manifold from high dimension space through new methods of dimensional reduction.

Shankar, M.G, Babu [1] explored a way for selecting classifiers using global and non-linear dimensionality reduction method. Tonglin Zhang and Baijian Yang [2] approaches an idea to obtain an array of adequate statistics by examining data in rows which provides exact solution for regression in Principal Component Analysis (PCA). Jing Wang and Chein-I Chan [3] discussed about dimensionality reduction using big data technologies and machine learning algorithm, where it uses DTML technique which features on multi-dimensional data reduction with high data reduction in duplicate data. In this paper, Sani M. Isa, Ito Wasito, Aniati Murni Arymurthi [4] proposed to classify sleep stage from ECG signal using Kernel Dimensionality Reduction (KDR), KDR is a supervised dimensionality reduction technique to maintain a relationship between input variables and output data. Monica Fira, Liviu Gorați [5] reviewed about dimensional data reduction using Laplacian Eigen map and Locality Preserving Projections especially for normal and pathological ECG signals. R.Sujatha, Sree Dharinya, G.Uma Maheswari [6] gives a comparative study and implementation on dimensionality reduction using fuzzy classifier for the purpose of disease diagnosis. Zhinan Li, Wenyao Xu, Anpeng Huang, and Majid Sarrafzadeh[7] provides introduction about manifold based approach for the detection of ECG anomaly usingdimensionality reduction method.
This paper proffers the clear illustration and uses of methods such as PCA, Kernel PCA and LLE. Principal Component Analysis is a prime linear method used for highly correlated data. PCA calculates the covariance matrix of input data and projects the original data to the new low dimensional data. Kernel PCA is linearly separable which uses a kernel function to view dataset into a higher dimensional feature space. Local Linear Embedded (LLE) is an efficient technique used to overcome non-linear dimension reduction problems.

1. EXISTING METHODOLOGY
Kernel Dimensionality Reduction Technique is used to compress the dimension of feature vector. This is helpful in accessing and comparing the performance of sleep stage classification before and after reducing the dimensional data. The block diagram of KDR Method for sleep stage classification is shown in the Figure.1.

![Figure 1. Block Diagram for Existing method](image)

KDR Method uses mainly five stages such as Feature Extraction, preprocessing, Dimensionality Reduction, Classification and Performance Evaluation.

The description for sleep stage classification using Kernel Dimensionality Reduction (KDR) Method is given below.

1.1. Preprocessing
Preprocessing is the process which uses QRS Detection time and RR Interval Selection where all the features in this KDR Method based on this Detection time. QRS Detection time is used to detect the time of occurrence of the QRS Complex in ECG Signal. RR Interval selection gives the interval between the peaks of two QRS complex.

1.2. Feature Extraction
Feature Extraction is a method based on epoch features and are processed based on QRS complex timing. For each epoch, HRV Features are extracted. In the stage of feature extraction, RR Interval between 0.5 and 1.5 is processed. RR Interval can also be normalized in order to remove subject dependencies.

1.3. Kernel Dimensionality Reduction
Kernel Dimensionality Reduction (KDR) approaches semi-parametric statistical framework.

\[
\min_{B \in \mathbb{R}^{p \times m}} \frac{\det S_{YU}}{\det S_{Y} \det S_{U}} \quad \text{Where} \quad Y = B^T X
\]  

(2.1)

The minimization problem referred with respect to the choice of matrix A or subspace S known to be as Kernel Dimensionality Reduction (KDR).
In order to determine effective dimensions, the original vector dimensions will be reduced to 2, 3 and 4 using KDR. The classification results are then compared to the classification results with a new reduced feature vector using an original feature vector.

### 1.4. Classification

Classification methods used for sleep stage classification are kNN as the baseline method, random forest and SVM. Random forest and SVM is mostly used because sleep stage data have characteristics of imbalanced dataset. Classification was performed on each epoch to determine the sleep stage of each epoch using two different scenarios such as subject independent classification and subject specific classification to find out the subject dependency factor.

Data for training and testing acquired from the same record in the independent classification of the subject but in the different classification of the subject are from the mixture of all documents.

### 1.5. Performance Evaluation

This stage uses two performance evaluation measurements known to be as accuracy and Cohen’s Kappa coefficient. Cohen’s kappa coefficient (κ) is a statistical measurement of inter-annotator agreement or inter-rater agreement for categorical items. Accuracy is also a statistical measurement which shows whether the condition is identified correctly or excluded.

### 2. PROPOSED METHODOLOGY

Classification problem in machine learning have factor called random variable, it gets harder to predict the training set as there are more number of features. Since, there is correlation and redundancy of high number of features dimensionality reduction algorithm plays a tremendous role. Dimensionality reduction is the process of compressing the dimension of the feature set. The block diagram of proposed system is shown in the Figure. 2.

![Block Diagram for Proposed method](image)

**Figure. 2.** Block Diagram for Proposed method

The description for Dimensionality Reduction using PCA, Kernel PCA and LLE is given below. The Source of input for the compression method is taken from the MIT-BIH Arrhythmia database.

#### 2.1. Principle Component Analysis (PCA)

Principle Component Analysis is an unsupervised linear (method or transformation) to decrease the variables of large data set into a smaller one and it is to be commonly known as PCA. PCA acts as a tool for making predictive models and in exploratory data analysis. PCA especially designed for unlabelled data and it also visualizes genetic distance. Principle Component can remove noise in the data. Firstly, this algorithm centres the data by subtracting the mean, it places the axis in the direction of large variation and then it extracts another axis orthogonal to the first axis and covers all other remaining variations as
possible. Covariance matrix will be diagonal when all the variation is along the coordinate axis. PCA is a method done by single value decomposition of a matrix or PCA can also be made by computing the covariance matrix of original data and executing the eigenvalue decomposition of calculated covariance value. Consider n-samples having k-variables to represent the samples.

\[
Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1k} \\ y_{21} & y_{22} & \cdots & y_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nk} \end{bmatrix} = [y_1, y_2, \ldots, y_k] \quad (3.1)
\]

Steps to analyse PCA are:

Step 1: Standardization

Standardization of dataset is achieved by

\[
y'_{ij} = \frac{y_{ij} - \bar{y}_j}{\sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (y_{ij} - \bar{y}_j)^2}} \quad (3.2)
\]

Since

\[
\bar{y}_j = \frac{1}{n} \sum_{i=1}^{n} y_{ij} \quad (3.3)
\]

Step 2: Computation of correlation matrix and corresponding Eigenvectors

Correlation matrix is given by

\[
C = (c_{ij})_{k \times k} \quad (3.4)
\]

and its corresponding eigenvalue \( \lambda_j \) is calculated using

\[
|\lambda E - C| = 0 \quad (3.5)
\]

Eigenvectors associated with these eigenvalues will be obtained.

2.2. Kernel PCA

Kernel PCA or KPCA is the reconstruction of linear PCA technique formed using kernel function. Kernel PCA uses Kernel ridge regression and Support vector machine, whereas in existing methodology KDR uses kNN, Random forest and SVM for sleep stage classification. KPCA is used to calculate the principle eigenvalues and eigenvectors of the kernel matrix. Kernel matrix constructed using kernel function is comparable to the inproducts of datapoints in the high dimensional dataset. When there is multiple hidden layers are present, auto-associators can be used. The non-linear mapping is constructed by kernel space in kernel PCA using the properties of PCA.

Kernel matrix \( K \) is calculated by

\[
K_{ij} = \kappa(y_i, y_j) \quad (3.6)
\]

\( \kappa \) is known kernel function and \( y_i \) is the datapoints in kernel matrix.

by modifying the kernel matrix \( K \), it becomes

\[
k_{p_{ij}} = k_{p_{ij}} - \frac{1}{n} \sum_{m} k_{p_{im}} + \frac{1}{n^2} \sum_{m,n} k_{p_{mn}} \quad (3.7)
\]

Mean of the feature is subtracted to make it as zero-mean in high dimensional space. Eigenvectors of the covariance matrix \( \beta_j \) is obtained by
The eigenvectors of the kernel matrix are given by
\[ \beta_l = \frac{1}{\sqrt{\lambda_l}} v_l \]  
(3.8)

where \( v_l \) is the eigenvector of the kernel matrix.

The covariance matrix in the kernel space is
\[ C = \frac{1}{N} \sum_{n=1}^{N} \Phi(y_n) \Phi(y_n)^T \]  
(3.9)

which produces eigenvector:
\[ V = \sum_{j=1}^{N} \beta_j \Phi(x_j) \]  
(3.10)

where \( \beta_j \) is the eigenvector of the ‘kernelized’ problem.

Low-dimensional data representation of KPCA is projected by
\[ Z = \left\{ \sum_j \beta_j \kappa(y_j y), \sum_j \beta_j \kappa(y_j y), \ldots, \sum_j \beta_j \kappa(y_j y) \right\} \]  
(3.11)

2.3. Local Linear Embedding (LLE)

Local Linear Embedding is a dimensionality reduction technique which is used to modify high-dimensional graphical data into low-dimensional graphical representation (similar to Isomap). LLE can construct data manifolds by making datapoints as linear combinations of its nearest neighbors. LLE can also easily retain weights in the linear combinations during the data represented in low dimension. Reconstruction error in LLE occurs when there is non-linearity in data and it is measured by calculating the distance between the original data and reconstructed data.

Assuming that the manifolds which are locally linear in the datapoint \( z_i \) in which datapoints formed as linear combinations \( W_i \) (and also called as reconstructed weight in low-dimensional space) of \( k \)-nearest neighbors \( z_i \). Since weights of the datapoints are invariant to rescaling, translation, and rotation so there will be linear mapping which will reconstruct weights. The reconstructed weight \( W_i \) will change the datapoint \( z_i \) by its neighbor in the high-dimensional representation which will also modify (or reconstruct) datapoint \( x_i \) by its neighbor in the low-dimensional representation.

3. RESULT AND DISCUSSION

This Section projects the empirical result of PCA, KPCA and LLE reduction techniques and gives the estimation of its performance. Figure 3(a) shows graphical representation of Input (ECG signal) which is having high-dimensional data from MIT-BIH Arrhythmia database. The Inputs from ECG signal is a source extracted from MIT-BIH records. These sources are taken as samples from the MIT-BIH database which is to be in the form of text converted to excel file. Then these samples in excel file is called in MATLAB where PCA, KPCA and LLE methods are applied for the data reduction. Dimensionality of Original dataset is 99.
Figure 3(a). High Dimensional ECG Input

Figure 3(b). Low Dimensional output using PCA

Figure 3(c). Low Dimensional output using KPCA
The low dimensional output which is got up by using the method called Principle component analysis is pictured in Figure.3(b). PCA is linear and optimal scheme used in applications such as data compression, image processing, face and pattern recognition, Visualisation etc.. Performance estimation has 82% of sensitivity and 89% of accuracy. PCA has low memory than KPCA. Sometimes there occurs loss of information. Dimensionality of original dataset is compressed to 25.

The compressed output of Kernel PCA is expressed in Figure.3(c). Kernel PCA have kernel space where it constructs the nonlinear mappings and it is also used in some of the applications such as speech recognition, face recognition and novelty detection. It has 85% Sensitivity and 92% accuracy. KPCA always prefers to handle the non-linear dataset. KPCA have dimensionality which is reduced to 20.

Figure.3(d). represents the output of Local Linear Embedding which is reduced dimensional data. LLE can easily retain weights by its neighboring datapoint in low dimension. Locally Linear Embedded is used in superresolution and sound source localization. Dimensionality of LLE is 18. Performance Analysis of LLE have estimated an accuracy of 93.02% and Sensitivity of 89%.

Performance Analysis is shown in Table.1

| Performance Analysis | PCA  | KPCA | LLE  |
|----------------------|------|------|------|
| **Sensitivity**      | 82   | 85   | 89   |
| **Accuracy**         | 89   | 92   | 93.02|

4. CONCLUSION

The Execution of ML algorithms with large input feature set can degrade its performance so there is a need for Dimensionality Reduction. This attempt has been made to compress such degradation problems. The proposed work deals with the theoretical and empirical view of linear and non-linear dimensionality reduction techniques such as Principle Component Analysis, Kernel PCA, Local Linear Embedded. Among the three reduction techniques KPCA can reduce computational complexity and LLE can easily overcome non-linear reduction problem. Performance Analysis of these linear and non-linear algorithm is estimated and compared.

References

[1] Shankar, M.G, Babu, C.G.; An exploration of ECG signal feature selection and classification using machine learning techniques; Int. J. Innovative Technol. Exploring Eng. Regul; Issue 3; Volume 9; 797–804; 2020.

[2] Tonglin Zhang and Baijian Yang; Big Data Dimension Reduction using PCA; IEEE International Conference on Smart Cloud; 152-157; 2016.

[3] Jing Wang and Chein-I Chan; Independent Component Analysis-Based Dimensionality Reduction With Applications in Hyperspectral Image Analysis; IEEE Transactions on Geoscience and Remote Sensing; VOL. 44; NO. 6; 1586-1599; 2006.

[4] Sani M. Isa, Ito Wasito, Aniati Murni Arymurthy; Kernel Dimensionality Reduction on Sleep Stage
Classification using ECG Signal; IJCSI International Journal of Computer Science Issues; Vol. 8; Issue 4; No 2; 115-123; 2011.

[5] Monica Fira, LiviuGoraú; Dimensionality Reduction for ECG Signals; Laplacian Eigenmaps and Locality Preserving Projections; IEEE Transactions; 2019.

[6] R. Sujatha, SreeDharinya, G. UmaMaheswari; Comparative Study on Dimensionality Reduction for Disease Diagnosis using Fuzzy Classifier; International Journal of Engineering and Technology; 79-84; 2018.

[7] Zhinan Li, WenYao Xu, Anpeng Huang, and Majid Sarrafzadeh; Dimensionality Reduction for Anomaly Detection in Electrocardiography; A Manifold Approach; Ninth International Conference on Wearable and Implantable Body Sensor Networks; 161-165; 2012.

[8] S. RangaSwamy, P S V SrinivasaRao, J. V N Raju, M. Nagavamsi; Dimensionality Reduction using Machine Learning and Big Data Technologies; International Journal of Innovative Technology and Exploring Engineering (IJITEE); Volume-9; Issue-2; 1740-1745; 2019.

[9] Chuan HE, Zhe DONG, Ruifan LI, Yixin ZHONG; Dimensionality Reduction for Text Using LLE; IEEE Transactions; 152-157; 2008.

[10] https://en.wikipedia.org/wiki/Dimensionality_reduction.