

**Department of Computer Science and Engineering(UG Studies)**

**PES University, Bangalore-560085**

|  |  |
| --- | --- |
| **Session :** Aug - Dec 2017  **Credits :** 0-0-2-0-1 | UE14CS405 : Machine Learning Lab |
| **Lab # : 10** | Implement PCA and estimate accuracy |

|  |
| --- |
| Lab 9:  Implement Principal Component Analysis |
| DataSet:  Iris data set  Given below is the 1st 10 rows of the Iris dataset where the columns are sepal length,sepal width, petal length, petal width. class attribute is not taken.  4.9,3.0,1.4,0.2  4.7,3.2,1.3,0.2  4.6,3.1,1.5,0.2  5.0,3.6,1.4,0.2  5.4,3.9,1.7,0.4  4.6,3.4,1.4,0.3  5.0,3.4,1.5,0.2  4.4,2.9,1.4,0.2  4.9,3.1,1.5,0.1  5.4,3.7,1.5,0.2 |
| Steps for finding Principal Components  1. **Normalize the data if needed to have features on the same scale since we are using Covarience matrix(Given dataset is already scaled)**  2. **Calculate the covariance matrix.**  **(You can use numpy package for Covarience matrix calculation)**  **Covariance is the measure of how two different variables change together. The covariance between two variables, X and Y, can be given by the following formula.**  *Cov(X, Y) = Σ ( Xi - mean(X) ) ( Yi -mean( Y) ) / N*  3. **Find the eigenvectors of the covariance matrix.**  **Note: Experiment 2 has been done on finding the eigen value and eigen vector**  4. **Translate the data to be in terms of the components**. |
| To Do list:Perform the above steps and fill the missing code attached |
| OUTCOME:  1)Dimentionality Reduction Technique  2)Used in Eigen face Recognition  3)Also used for Classification problems  4)can form cluster of same type |