



California State University Long Beach

BME 455 Spring 2021

Lab 9

Instructor: Elsa Harris

In this lab you will work with Principal Component Analysis (PCA).

Objectives:

1. Use sklearn to visualize a trained PCA analysis
2. Train any ML model using PCA in scikit-learn
3. Compare ML model without and with PCA to evaluate the performance.

Part 1: Complete the work from “PCA using Python (scikit-learn) by Michael Galarnyk:

<https://towardsdatascience.com/pca-using-python-scikit-learn-e653f8989e60>

Part 2: This part will be a bit different from what you have been doing previously. It is going to be free style, which means you can search on google, Knovels, or any other resources to complete all tasks.

Datasets: Use PIMA Diabetes Datasets (From old lab)

Taks1(Load data): Load the Diabetes datasets, output the data frame.

Task2(Create Datasets): Create the training and testing data. Show that you successfully split those data. You can decide the proportion of splitting. Use all features of the data! (ex:80% for training and 20% for testing)

Hint: create X\_train,X\_test,Y\_train,Y\_test and so on...

Task3(Modeling): Choose one ML model that you think is the best to approach for this lab

Task4(PCA): Apply PCA techniques to your datasets, set your principal component to one. (Make sure to utilize “pca.fit\_transform” to normalize your datasets before processing PCA and try many values in principle components to find out the best result. Hint: choose n\_components = between 0 and 1 for diabetes datasets)

Task4(Results): Perform prediction for the model you choose, and output both results from without PCA and using PCA.

Task4(n\_components): Set n\_components =2 or 3 to see the result

Task6:(Report): Write a report of results after making the comparison of results. Write the main purpose of using PCA, pro and con of using PCA, when to use PCA, and state the reason why you chose the specific ML model, things you learned from applying PCA technique in this lab.

Name each submission file as Lab7\_PartX\_FirstName\_LasstName