

**Lab 3 report**

**Submitted to:**

Dr. Raihan Ul Islam

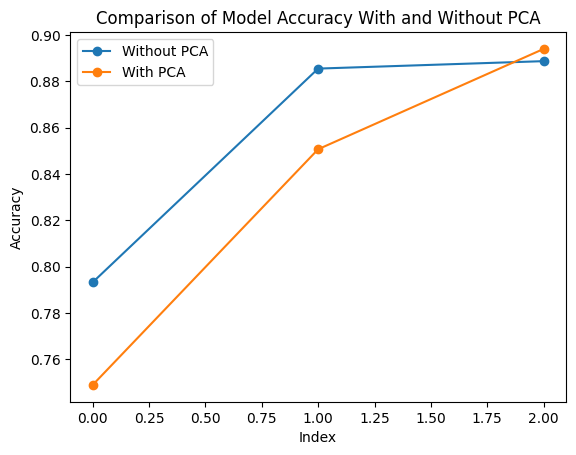
**Submitted by:**

Md. Arnob Tanvir

**ID:** 2020-1-60-220

**Date of submission:** 07/24/2024

The results are summerized in the following picture.



First one is Decision tree, 2nd one is random forest and the last one is SVM. Here’s an analysis

**Without PCA:**

* + **Decision Tree**:0.793 (Highest accuracy)
  + **Random Forest**: 0.885 (Moderate accuracy)
  + **SVC**: 0.888 (Similar accuracy as Random Forest, slightly lower)

**With PCA:**

* + **Decision Tree**: 0.75 (Significant drop in accuracy)
  + **Random Forest**: 0.85 (A small drop in accuracy, but still higher than the Decision Tree with PCA)
  + **SVC**: 0.894 (Significant improvement, surpassing the Decision Tree and close to Random Forest without PCA)

#### ****Analysis****

**Decision Tree**:

* + **Without PCA**: The Decision Tree performs best without PCA because it utilizes all the original features to make decisions, leading to high accuracy.
  + **With PCA**: The accuracy drops because PCA reduces the feature space, possibly removing some features that the Decision Tree would have used to make accurate splits.

**Random Forest**:

* + **Without PCA**: The Random Forest performs well due to its ensemble nature, which reduces overfitting and improves generalization.
  + **With PCA**: The slight drop in accuracy indicates that the ensemble method still benefits from more features but can handle reduced dimensions better than a single Decision Tree.

**Support Vector Classifier (SVC)**:

* + **Without PCA**: The SVC performs comparably to the Random Forest but slightly worse, as it struggles with high-dimensional data without prior dimensionality reduction.
  + **With PCA**: SVC sees a noticeable increase in accuracy, suggesting that PCA effectively reduces the dimensionality and noise, making it easier for the SVC to find a clear separation between classes.

#### 5. ****Conclusion****

**PCA** has a mixed impact on different models. While it improves the performance of SVC by reducing the dimensionality and simplifying the decision boundary, it negatively impacts the Decision Tree by removing potentially useful features. Random Forest, being more robust, only slightly suffers from PCA but still maintains a high level of accuracy.

**Recommendation**: When using PCA, it is essential to consider the nature of the model being applied. For instance, PCA is highly beneficial for SVC but not as much for Decision Trees. Therefore, model selection should be informed by how well the model can handle reduced feature spaces.

This analysis highlights the importance of model selection and preprocessing techniques like PCA in achieving optimal performance on complex datasets like Fashion MNIST.