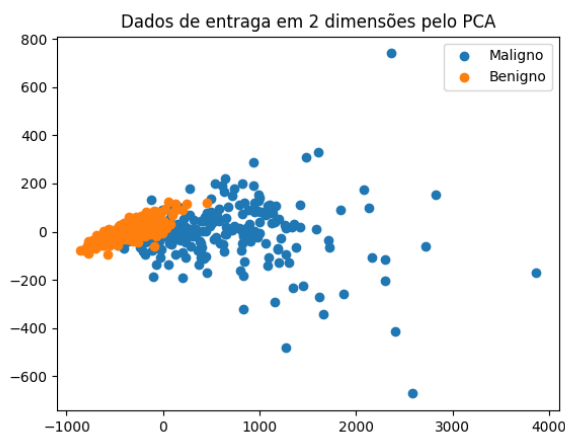

RELATÓRIO PCA – SVM

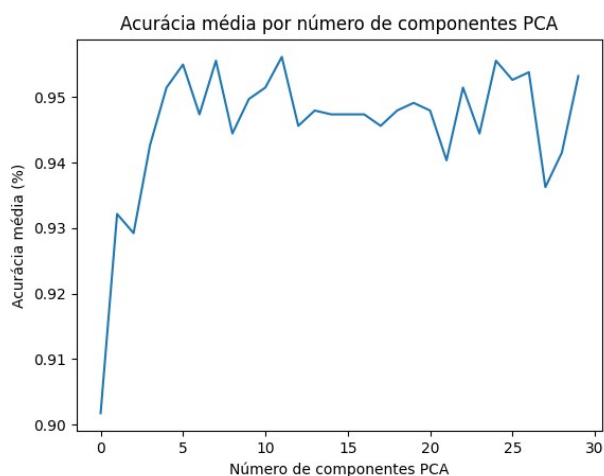
GABRIEL SARAIVA ESPESCHIT – 2015065541

04 de out de 2020

Nesse exercício, iremo implementamos a técnica de PCA para reduzir as dimensões dos dados do conjunto *Breast Cancer*. Em seguida utilizamos o SVM para fazer a classificação desses mesmos dados. Para as duas técnicas, utilizou de funções do módulo *Sklearn*. Os dados projetados em 2 dimensões podem ser visualizados abaixo.



Fizemos uma rotina para verificar quantas dimensões seria adequado usar para fazer a classificação dos dados. O gráfico abaixo mostra a acurácia contra o número de dimensões que usamos para os dados de entrada. Para cada número de dimensões, fizemos 30 iterações para um conjunto dividido em 50% de teste e 50% de treino.



Como pode ser observado, a melhor acurácia foi obtida com 12 dimensões. Sendo assim, usou esse valor e, aplicando a técnica de K-Folds, fez-se o processo de treino e teste para 10 *folds*. Os resultados obtidos encontram-se abaixo.

Acurácia média: 0.94

Acurácia por fold: [0.912, 0.982, 0.982, 0.929, 0.964, 0.929, 0.929,
0.964, 0.929, 0.910]