



Visualizing PCA

Profa. Dra. Roseli Aparecida Francelin Romero

Disciplina: SCC 0270 - Redes Neurais

Lucas Yudi Sugi - 9293251

Introdução

PCA é uma das técnicas mais usadas atualmente para realizar a redução de dimensionalidade, permitindo que haja uma melhor visualização dos dados assim como uma melhora nos algoritmos de aprendizado de máquina. Conhecê-lo é fundamental para se trabalhar com ciência de dados.

Objetivo

Visualizar o novo espaço de dados checando se houve melhoras.

PCA

Pelo gráfico abaixo é possível verificar que existem atributos “bagunçando” as classes como ocorre em pc2 vs pc4. Contudo, é interessante notar como os componentes pc1 e pc2 conseguem separar bem as classes, além de permitir uma boa visualização dos dados.

Note que poderíamos utilizar apenas ambos para realizar uma análise exploratória, assim como utilizá-los para classificação.

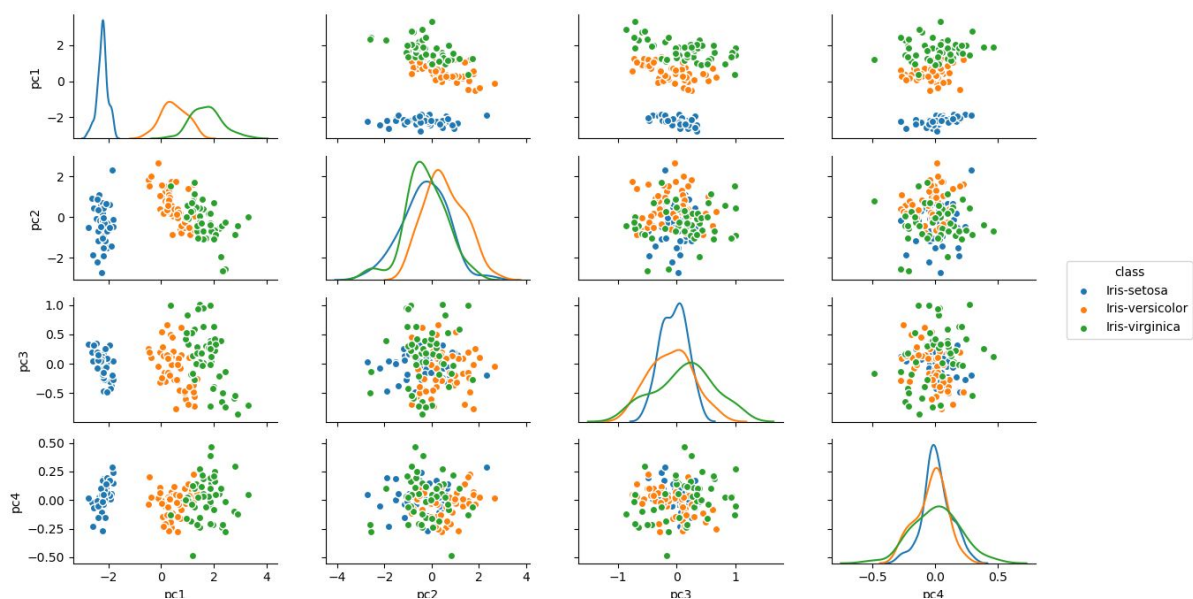


Figura 1: Scatter plot do novo espaço gerado.

Conclusão

A utilização do PCA permitiu uma melhor separação no espaço de dados, de modo que com menos atributos alcançamos as mesmas informações do dataset original.