

PCA:

1. Model bekerja dengan menghitung matriks kovarians dari data yang diberikan. Matriks ini adalah ukuran korelasi antara satu variabel dengan variabel lainnya dalam suatu dataset. Kemudian, akan dihitung nilai eigen dan vektor eigen dari matriks kovarians. Nilai eigen menjelaskan seberapa besar variansi dari dataset yang dijelaskan oleh masing-masing komponen utama. Nilai eigen akan diurutkan dari yang terbesar ke yang terkecil. Vektor eigen dari nilai eigen yang tertinggi akan dipilih untuk menjadi komponen utama. Kemudian, data diubah ke dimensi yang lebih rendah dengan melakukan *dot product* antara data yang diberikan dengan vektor eigen yang telah dipilih sebelumnya.
2. Hasil yang diperoleh antara model bawaan dari scikit learn dengan model *from scratch* menunjukkan hasil *explained variance* yang sama. Namun, Ketika diplot, hasil plotnya menunjukkan ada perbedaan, hal ini disebabkan oleh orientasi data yang berbeda yang bergantung pada tanda vektor eigen yang terpilih. Vektor eigen bisa saja berbeda tanda sehingga menyebabkan data diplot pada sumbu yang sama, tetapi dengan orientasi yang berbeda.