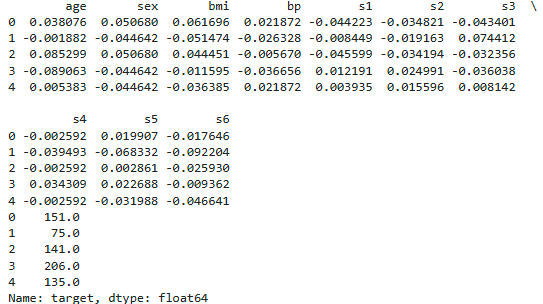
**KELOMPOK:**

Moh. Naufal Faqih 10222044

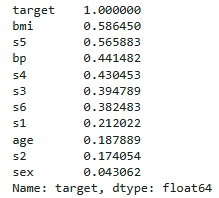
Firman Firdaus 10222033

Ryan Azis S. 10222041

1. **Fitur Seleksi**
   1. Memuat Dataset Diabetes



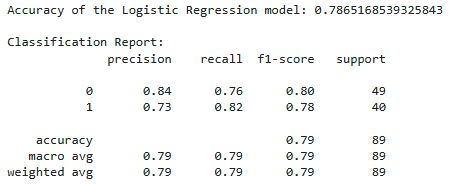
* 1. Menghitung Korelasi Pearson antara Fitur dengan label



* 1. Memilih 5 Fitur Terbaik Menggunakan RFE



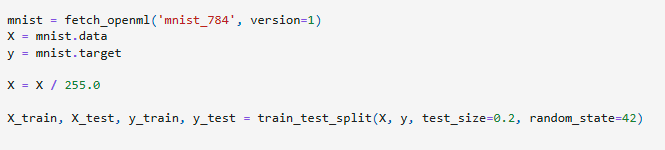
* 1. Membuat Model Prediksi & Evaluasi Performa Model



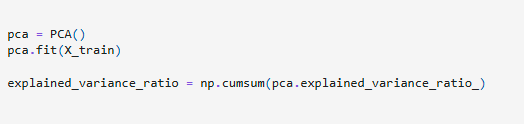
Hasil:

* Konversi label ke bentuk biner membuat model klasifikasi jauh lebih efektif dibandingkan model regresi awal.
* Akurasi model sebesar ~79% menunjukkan bahwa model cukup baik dalam membedakan pasien dengan risiko tinggi dan rendah.
* Precision dan recall seimbang, baik pada kelas 0 maupun kelas 1, menunjukkan bahwa model tidak bias secara signifikan terhadap salah satu kelas.
* Macro dan weighted average f1-score = 0.79 menandakan performa stabil di kedua kelas.

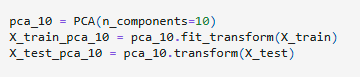
1. **Reduksi Dimensi**
   1. Muat MNIST Dataset



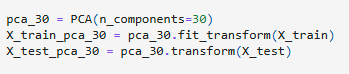
* 1. Gunakan PCA untuk Mereduksi Dimensi



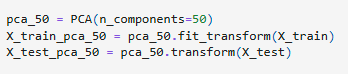
* 1. PCA dengan 10 Components



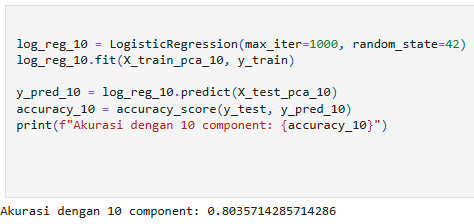
* 1. PCA dengan 30 Components



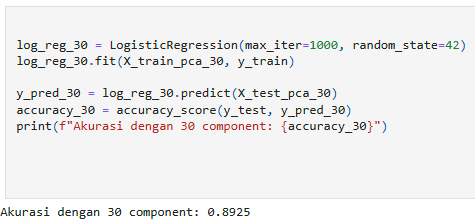
* 1. PCA dengan 50 Components



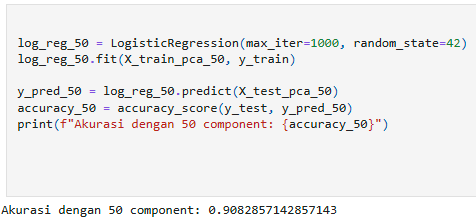
* 1. Latih Model Pada 10 Components



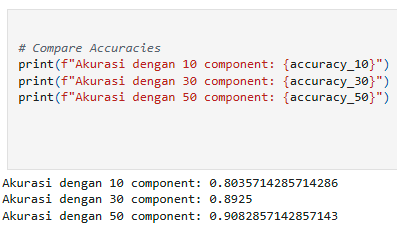
* 1. Latih Model pada 30 Component



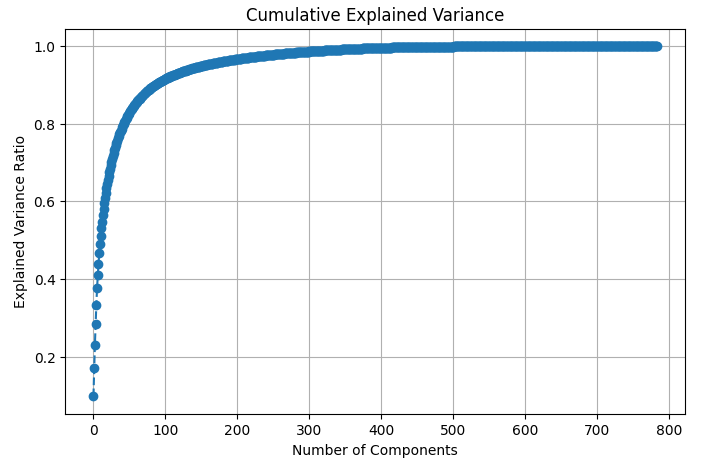
* 1. Latih Model pada 50 Component



* 1. Bandingkan Model Akurasi



* 1. Visualisasi Total Variansi



Hasil:

Untuk menjaga keseimbangan antara akurasi dan efisiensi, 30 komponen adalah pilihan yang tepat karena memberikan peningkatan akurasi yang signifikan dari 10 komponen (+8.89%), memiliki kompleksitas yang lebih ringan dibandingkan 50 komponen, namun hanya kehilangan sedikit akurasi (sekitar 1.6%).