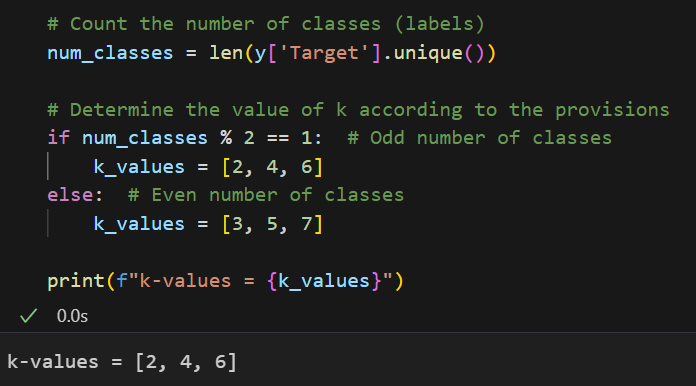
Nama: Nabhan Nabilah

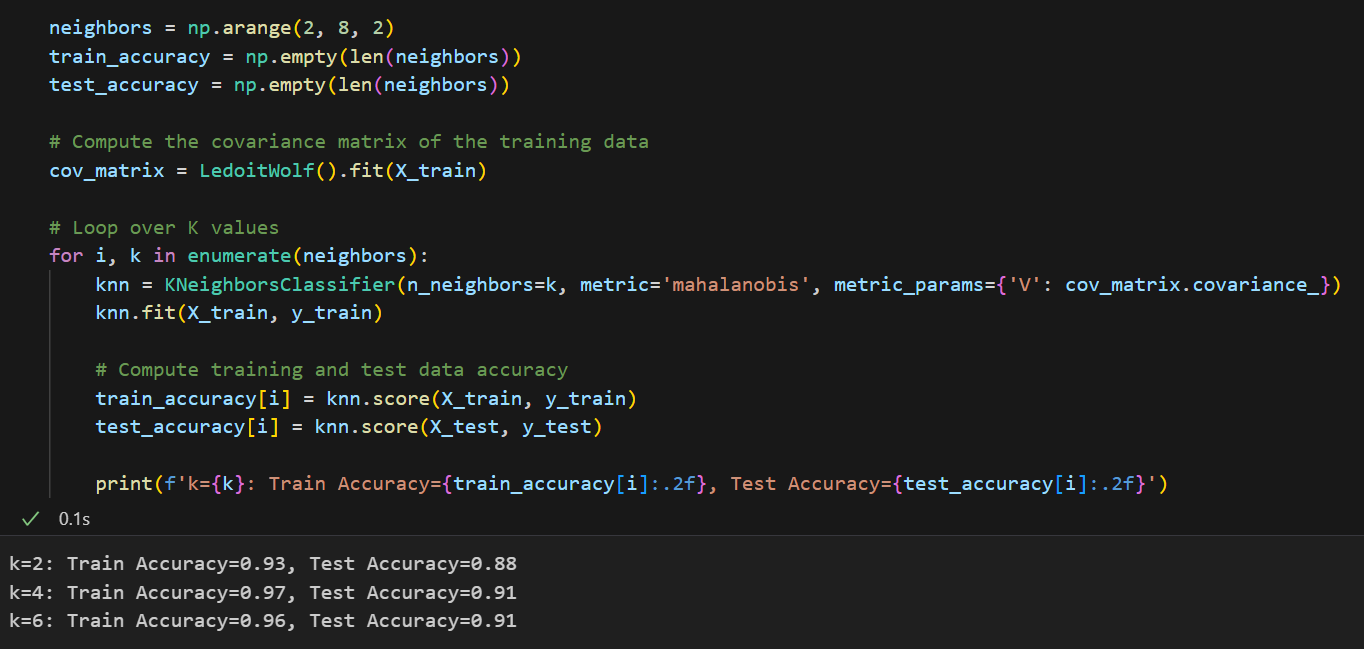
NIM: 4112321013

**Hasil Klasifikasi dengan K-Nearest Neighbors (KNN)**

Dalam analisis KNN ini, kita mempertimbangkan nilai k yang optimal untuk klasifikasi anggur. Pengambilan keputusan terkait nilai k didasarkan pada jumlah kelas pada kolom target. Karena jumlah kelas pada kolom target adalah bilangan ganjil (3 kelas), kita memilih nilai k yang juga ganjil.

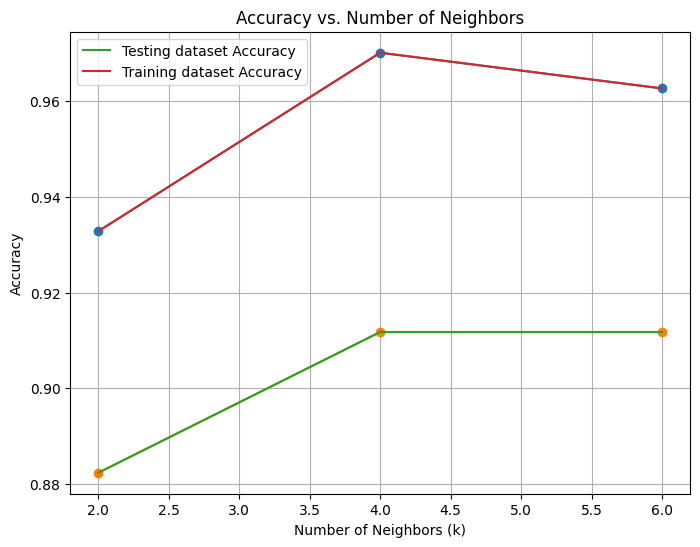


Proses berikutnya melibatkan pembagian dataset menjadi dua bagian, yakni data pelatihan (80%) dan data pengujian (20%). Model KNN dilatih menggunakan metode jarak Mahalanobis dengan berbagai nilai k, yaitu 2, 4, dan 6.

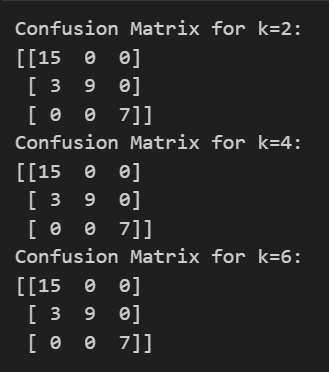


Hasil evaluasi menunjukkan bahwa k = 4 memberikan akurasi tertinggi pada data pengujian, yakni sebesar 91%. Walaupun akurasi pada k = 6 juga mencapai angka yang sama, seringkali lebih baik memilih model yang lebih sederhana ketika ada dua model dengan akurasi yang setara. Oleh karena itu, nilai k yang dipilih adalah k = 4.

Berikut adalah visualisasi data pelatihan dan pengujian berdasarkan nilai k:



Selain itu, berikut adalah matriks kebingungan (confusion matrix) yang menggambarkan performa model KNN pada k = 4:



Dalam keseluruhan analisis ini, k=4 adalah pilihan terbaik karena menghasilkan akurasi pengujian tertinggi dibandingkan dengan nilai k lainnya. Selain itu, pemilihan model yang lebih sederhana dengan akurasi setara adalah praktek yang baik dalam pemodelan.