

Nama	: Adhimas Aryo Bimo
NIM	: 13523052
Model	: KNN

1. Sederhananya, KNN merupakan algoritma supervised berbasis instance (lazy learning). Prinsipnya adalah memprediksi label suatu data uji, algoritma ini akan mencari k tetangga terdekat dari data train lalu menentukan label mayoritas untuk klasifikasi atau rata-rata untuk regresi.

Dalam mengetahui jarak tiap label, KNN akan menggunakan fungsi jarak yang beragam. Dalam hal ini dapat digunakan fungsi jarak manhattan, minkowski, atau euclidian.

Karena basisnya yang merupakan lazy learning. Dalam melakukan fit, KNN tidak melakukan training kompleks, melainkan hanya menyimpan data latih.

Hal penting dalam KNN adalah ketika melakukan predict. Prediksi dilakukan dengan menghitung jarak data uji terhadap semua data latih. Hasil prediksi ditentukan berdasarkan label mayoritas untuk klasifikasi atau median untuk regresi.

2. KNN yang saya buat dengan KNN yang berada di scikit learn memiliki skor yang mirip, dapat dilihat pada tabel berikut :

===== KNN_custom =====	===== KNN_sklearn =====
Fold 1: 895.196	Fold 1: 895.196
Fold 2: 1010.578	Fold 2: 1010.578
Fold 3: 944.400	Fold 3: 944.400
Fold 4: 954.468	Fold 4: 954.468
Fold 5: 976.441	Fold 5: 976.441
Mean RMSE: 956.216 ± 38.027	Mean RMSE: 956.216 ± 38.027

Mungkin karena modelnya sederhana dan memang tergolong lazy model jadi algoritma nya hanya sekedar mengukur jarak dari data prediksi dengan data yang di fit di train sehingga bisa memiliki nilai yang sama persis.

3. Model ini sangat lazy jadi agak sulit diimprove performanya untuk melakukan prediksi. Mungkin dengan menambah neighbors pada parameter model dapat meningkatkan performa walaupun tidak banyak.