

Tabel Perbandingan Metrik

Metric	k-NN	Decision Tree	Keterangan
Accuracy	0.8385	0.7815	k-NN lebih tinggi
Precision	0.6848	0.5279	k-NN lebih presisi
Recall	0.6097	0.8756	Decision Tree lebih baik deteksi >50K
F1-Score	0.6451	0.6587	Decision Tree sedikit unggul
ROC AUC	0.8705	0.8917	Decision Tree lebih baik bedakan kelas

Analisis Per Metrik

1. Accuracy (Akurasi)

- k-NN lebih unggul (83.85% vs 78.15%)
- *Tapi kurang relevan jika data tidak seimbang* (bias ke kelas mayoritas $\leq 50K$)

2. Precision

- k-NN jauh lebih baik (68.48% vs 52.79%)
- Artinya: Dari semua yang diprediksi >50K oleh k-NN, 68.48% benar (Decision Tree hanya 52.79%)
- *Krusial untuk minimalkan False Positive* (misal: hindari salah beri promosi ke yang income rendah)

3. Recall

- Decision Tree unggul signifikan (87.56% vs 60.97%)
- Artinya: Decision Tree bisa deteksi 87.56% orang berincome >50K yang ada di data
- *Krusial untuk minimalkan False Negative* (misal: jangan sampai lewatkan calon customer kaya)

4. F1-Score

- Decision Tree sedikit lebih baik (0.6587 vs 0.6451)
- *Balance antara Precision dan Recall*

5. ROC AUC

- Decision Tree lebih baik (0.8917 vs 0.8705)
- Artinya: Kemampuan membedakan kelas >50K dan ≤50K lebih baik

Kesimpulan

Pemilihan model terbaik tergantung dari kebutuhan. Jika ingin minimalkan kesalahan prediksi income tinggi (False Positive) Pilih k-NN karena Precision lebih tinggi (68.48%). Alasannya adalah lebih pasti saat memprediksi >50K (lebih sedikit orang miskin yang salah dikira kaya). Jika ingin maksimalkan deteksi income tinggi (minimalkan False Negative) Pilih Decision Tree karena Recall lebih tinggi (87.56%) alasannya adalah hanya 12.44% orang kaya yang terlewat (vs k-NN yang miss 39.03%). Jika butuh keseimbangan (F1-Score) dan kemampuan umum (AUC) pilih Decision Tree karena F1-Score lebih tinggi dan ROC AUC lebih baik