

1. Berikan pengertian *confusion matrix*

JAWAB

Confusion Matrix adalah tabel dengan 4 kombinasi berbeda dari nilai prediksi dan nilai aktual. Ada empat istilah yang merupakan representasi hasil proses klasifikasi pada confusion matrix yaitu True Positif, True Negatif, False Positif, dan False Negatif.

- True Positive (TP) :
Interpretasi: Anda memprediksi positif dan itu benar.
- True Negative (TN):
Interpretasi: Anda memprediksi negatif dan itu benar.
- False Positive (FP): (Kesalahan Tipe 1)
Interpretasi: Anda memprediksi positif dan itu salah.
- False Negative (FN): (Kesalahan Tipe 2, kesalahan tipe 2 ini sangat berbahaya)
Interpretasi: Anda memprediksi negatif dan itu salah.

2. Bagaimana metrik akurasi, *precision*, *recall* dan *F1 measure* (disebut juga *F1 score*) didapatkan dari *confusion matrix*. Lihatlah artikel berikut untuk menjawab:

<https://medium.com/@ksnugroho/confusion-matrix-untuk-evaluasi-model-pada-unsupervised-machine-learning-bc4b1ae9ae3f>

JAWAB

Kita dapat menggunakan confusion matrix untuk menghitung berbagai performance metrics untuk mengukur kinerja model yang telah dibuat. Pada bagian ini mari kita pahami beberapa performance metrics populer yang umum dan sering digunakan: accuracy, precision, dan recall.

Accuracy

Accuracy menggambarkan seberapa akurat model dapat mengklasifikasikan dengan benar. Maka, accuracy merupakan rasio prediksi benar (positif dan negatif) dengan keseluruhan data. Dengan kata lain, accuracy merupakan tingkat kedekatan nilai prediksi dengan nilai aktual (sebenarnya). Nilai accuracy dapat diperoleh dengan persamaan (1).

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision (Positive Predictive Value)

Precision menggambarkan tingkat keakuratan antara data yang diminta dengan hasil prediksi yang diberikan oleh model. Maka, precision merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Dari semua kelas positif yang telah di prediksi dengan benar, berapa banyak data yang benar-benar positif. Nilai precision dapat diperoleh dengan persamaan (2).

$$precision = \frac{TP}{TP + FP} \quad (2)$$

Recall atau Sensitivity (True Positive Rate)

Recall menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi. Maka, recall merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Nilai recall dapat diperoleh dengan persamaan (3).

$$recall = \frac{TP}{TP + FN} \quad (3)$$