

## Confusion Matrix

*Machine Learning* merupakan salah satu cabang dari disiplin ilmu *artificial intelligence* yang membahas bagaimana sistem dibangun berdasarkan pada data. Jadi *machine learning* merupakan proses komputer untuk belajar dari data (*learn from data*). Jika tidak ada data, komputer tidak akan bisa belajar. Salah satu teknik aplikasi pada *machine learning* adalah *supervised learning*. Klasifikasi merupakan *supervised learning*, yang merupakan model prediksi dimana hasil prediksinya bersifat diskrit.

Bagaimana mengukur performa dari model klasifikasi yang digunakan? Jawaban sederhana nya adalah membandingkan nilai aktual dengan nilai prediksi. *Confusion Matrix* adalah pengukuran performa untuk masalah klasifikasi *machine learning* dimana keluaran dapat berupa dua kelas atau lebih.

*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. Mari kita pahami apa itu True Positif, False Positif, False Negatif, dan True Negatif dapat dilihat pada Tabel 1.

Tabel 1. Confusion Matrix 2 Dimensi (2 kelas)

<b>True Positif (TP)</b>	<b>False Positif (FP)</b>
<b>False Negatif (FN)</b>	<b>True Negatif (TN)</b>

### Contoh kasus: prediksi seorang sakit jantung

True Positive (TP) :

Interpretasi: Anda memprediksi positif (sakit jantung) dan itu benar.

Anda memprediksikan bahwa seseorang sakit jantung dan memang benar dia sakit jantung.

True Negative (TN):

Interpretasi: Anda memprediksi negatif (tidak sakit jantung) dan itu benar.

Anda memprediksikan bahwa seseorang tidak sakit jantung dan benar dia tidak sakit jantung.

False Positive (FP): (Kesalahan Tipe 1)

Interpretasi: Anda memprediksi positif (sakit jantung) dan itu salah.

Anda memprediksikan bahwa seseorang sakit jantung tetapi dia tidak sakit jantung

False Negative (FN): (Kesalahan Tipe 2, kesalahan tipe 2 ini sangat berbahaya)

Interpretasi: Anda memprediksi negatif (tidak sakit jantung) dan itu salah.

Anda memperkirakan bahwa seseorang tidak sakit jantung tetapi sebenarnya dia sakit jantung.

Dari contoh di atas dapat digambarkan bahwa:

Nilai Prediksi adalah keluaran dari program dimana nilainya Positif dan Negatif

Nilai Aktual adalah nilai sebenarnya dimana nilainya True dan False

### **Contoh Confusion Matrix :**

Suatu perusahaan membuat sebuah model yang dilatih untuk memprediksi apakah seorang karyawan di perusahaan tsb terkena covid-19 atau tidak. Dengan asumsi perusahaan tsb mempunyai 175 karyawan. Dari model *classifier* memprediksi karyawan positif covid-19 sebanyak 145 dan karyawan negatif covid-19 sebanyak 30 tetapi pada kenyataannya, karyawan positif covid-19 sebanyak 150 dan karyawan negatif covid-19 sebanyak 25 (dapat dilihat pada Tabel 1).

Tabel 2 Contoh Confusion Matrix

<b>n= 175</b>	<b>Aktual : Positif (1)</b>	<b>Aktual : Negatif (0)</b>
<b>Prediksi : Positif (1)</b>	TP: 125	FP: 20
<b>Prediksi : Negatif (0)</b>	FN: 25	TN: 5
	150	25

Pada kasus ini :

True Positive (TP): kita memprediksi karyawan positif covid-19 dan memang benar karyawan tsb positif covid-19

True Negative (TN): kita memprediksi karyawan negative covid-19 dan memang benar karyawan tsb negatif covid-19.

False Positive (FP):kita memprediksi karyawan positif covid-19 dan ternyata prediksi salah, ternyata karyawan negatif covid-19

False Negative (FN) : kita memprediksi karyawan negatif covid-19 dan ternyata prediksi salah, ternyata karyawan tsb positif covid-19.

Seperti telah dijelaskan di atas bahwa FN merupakan kesalahan tipe 2 dimana kesalahan ini sangat berbahaya. Contoh : karyawan di prediksi negatif covid-19 padahal ternyata karyawan positif covid-19 maka karyawan tersebut terlambat mengetahui keadaan sebenarnya sehingga tidak segera dilakukan tindakan pencegahan pengobatan dan isolasi mandiri. Dimana karyawan dapat menularkan virus covid19 ke banyak karyawan lainnya dalam perusahaan dan dapat menyebabkan kematian untuk karyawan lain juga karyawan tsb, maka dapat dihitung nilai *accuracy*, *precision*, *recall* dan F-1 score.

*Accuracy* menngambarkan seberapa akurat model dalam mengklasifikasikan dengan benar

$$\begin{aligned} \text{Accuracy} &= (TP+TN) / (TP+FP+FN+TN) \\ &= (125+5) / (125+20+25+5) \\ &= 0.742 \\ &= 0.742 * 100 \% = 74.2 \% \end{aligned}$$

*Precision* menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang diberikan oleh model.  $Precision = (TP) / (TP + FP)$

$$\begin{aligned} &= 125 / (125 + 20) \\ &= 0.86 \\ &= 0.86 * 100\% = 86\% \end{aligned}$$

*Recall* atau *sensitivity*: menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi.  $Recall = TP / (TP + FN)$

$$\begin{aligned} &= 125 / (125+25) \\ &= 0.83 \\ &= 0.83 * 100\% = 83\% \end{aligned}$$

F-1 Score menggambarkan perbandingan rata-rata precision dan recall yang dibobotkan. Accuracy tepat kita gunakan sebagai acuan performansi algoritma jika dataset kita memiliki jumlah data False Negatif dan False Positif yang sangat mendekati (*symmetric*). Namun jika jumlahnya tidak mendekati, maka sebaiknya kita menggunakan F1 Score sebagai acuan.

$$F-1 \text{ Score} = (2 * \textit{Recall} * \textit{Precision}) / (\textit{Recall} + \textit{Precision})$$

$$= (2 * 0.83 * 0.86) / (0.83 + 0.86)$$

$$= 1.4276 / 1.69$$

$$= 0.84 * 100\%$$

$$= 84\%$$