

Nama : Akmal Zuhdy Prasetya
 NIM : H071191035
 Kelas : Pengantar Data Mining

Resume Pertemuan 8

PERFORMANCE MEASURE

A. Classification

Berikut merupakan hal-hal yang perlu diperhatikan serta beberapa cara mengukur performa dari suatu model klasifikasi dalam data mining.

1. Confusion Matrix

Di bidang machine learning dan khususnya masalah klasifikasi statistik, confusion matrix, juga dikenal sebagai error matrix, adalah tata letak tabel khusus yang memungkinkan visualisasi kinerja suatu algoritma, biasanya termasuk dalam supervised learning. Setiap baris matriks mewakili instance di kelas aktual sementara setiap kolom mewakili instance di kelas yang diprediksi, atau sebaliknya. Namanya berasal dari fakta bahwa memudahkan untuk melihat apakah sistem ini membingungkan dua kelas. Ini adalah jenis tabel kontingensi khusus, dengan dua dimensi, dan set "kelas" yang identik di kedua dimensi.

		Actual class		
		Positive	Negative	
Predicted class	Positive	TP: True Positive	FP: False Positive (Type I Error)	Precision: $\frac{TP}{TP + FP}$
	Negative	FN: False Negative (Type II Error)	TN: True Negative	Negative Predictive Value: $\frac{TN}{TN + FN}$
		Recall or Sensitivity: $\frac{TP}{TP + FN}$	Specificity: $\frac{TN}{TN + FP}$	Accuracy: $\frac{TP + TN}{TP + TN + FP + FN}$

Misalnya kita ambil contoh seorang pasien yang berobat ke dokter dengan gejala tertentu. Karena ini musim Covid, anggap saja dia demam, batuk, sakit tenggorokan, dan pilek. Ini adalah gejala yang dapat terjadi selama perubahan musim juga. Oleh karena itu, sulit bagi dokter untuk melakukan diagnosis yang tepat.

Adapun beberapa keterangan yang terdapat di dalam confusion matrix berdasarkan contoh kasus itu adalah sebagai berikut.

- True Positive (TP)

Katakanlah pasien tersebut benar-benar menderita Covid dan pada saat dilakukan asesmen yang diperlukan, dokter mengklasifikasikannya sebagai pasien Covid. Ini disebut TP atau True Positive. Hal ini karena kasusnya positif secara nyata dan sekaligus kasus tersebut diklasifikasikan dengan benar. Kini, pasien dapat diberikan pengobatan yang tepat yang artinya, keputusan yang diambil oleh dokter akan berdampak positif bagi pasien dan masyarakat.

- False Positive (FP)

Katakanlah pasien itu tidak menderita Covid dan dia hanya menunjukkan gejala flu musiman tetapi dokter mendiagnosisnya dengan Covid. Ini disebut FP atau False Positive. Ini karena kasus itu sebenarnya negatif tetapi salah diklasifikasikan sebagai positif. Sekarang, pasien akan dirawat di rumah sakit atau di rumah dan akan diberikan perawatan untuk Covid. Ini adalah ketidaknyamanan yang tidak perlu baginya dan orang lain karena dia akan mendapatkan perawatan dan karantina yang tidak diinginkan. Ini disebut sebagai Type I Error.

- True Negative (TN)

Katakanlah pasien tidak menderita Covid dan dokter juga memberinya chit bersih. Ini disebut TN atau True Negative. Hal ini karena kasus tersebut sebenarnya negatif dan juga tergolong negatif yang merupakan hal yang benar untuk dilakukan. Sekarang pasien akan mendapatkan pengobatan untuk penyakitnya yang sebenarnya daripada mengambil pengobatan Covid.

- False Negative (FN)

Katakanlah pasien itu menderita Covid dan dokter tidak mendiagnosisnya dengan Covid. Ini disebut FN atau False Negative karena kasusnya sebenarnya positif tetapi salah diklasifikasikan sebagai negatif. Sekarang pasien tidak akan mendapatkan pengobatan yang tepat dan juga akan menularkan penyakitnya kepada orang lain. Ini adalah situasi yang sangat berbahaya dalam contoh ini. Ini juga disebut Type II Error.

2. Accuracy

Akurasi adalah kedekatan pengukuran dengan nilai tertentu. Dalam istilah yang lebih sederhana, jika kita mengukur sesuatu berulang kali, kita mengatakan pengukuran itu akurat jika mendekati nilai sebenarnya dari kuantitas yang diukur.

Berikut adalah cara menghitung akurasi dari suatu model klasifikasi.

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} = \frac{TP + TN}{TP + FP + TN + FN}$$

3. Recall (True Positive Rate)

Recall mengukur fraksi contoh positif yang diprediksi dengan benar oleh classifier. Cara menghitung Recall adalah sebagai berikut.

$$Recall = Sensitivity = TP Rate = \frac{TP}{TP + FN}$$

4. Precision

Presisi adalah sejauh mana pengukuran berulang di bawah kondisi yang sama menunjukkan hasil yang sama. Hal ini sering diukur dengan standar deviasi dari satu set nilai.

Berikut adalah cara menghitung presisi dari suatu model klasifikasi.

$$Precision = \text{Positive Prediction Value} = \frac{TP}{TP + FP}$$

Precision menentukan pecahan record yang benar-benar menjadi positif pada grup yang telah dideklarasikan oleh classifier sebagai kelas positif. Semakin tinggi presisi, semakin rendah jumlah False Positive atau FP yang dilakukan oleh model.

5. F1-Score / F-Measure

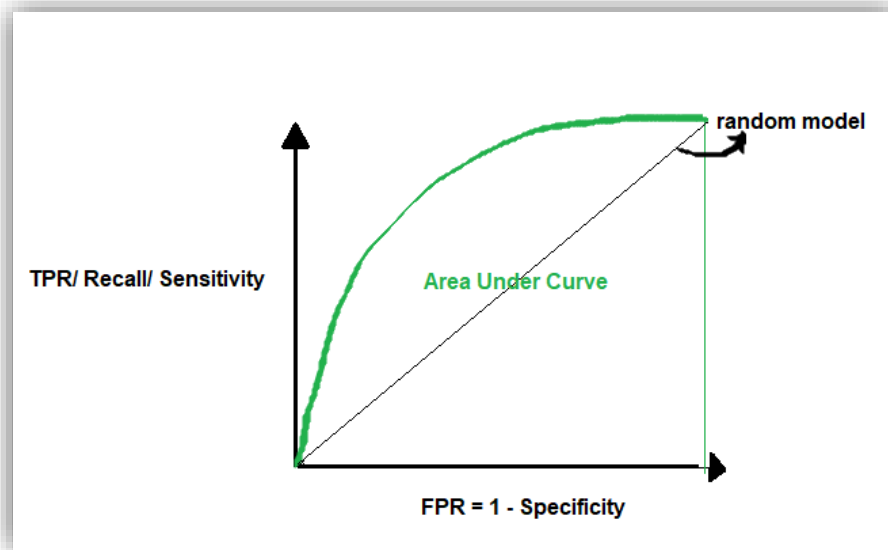
Dua metrik di atas yaitu presisi dan recall dapat digabungkan menjadi satu metrik yang disebut F1-Score atau F-Measure. Ini adalah rata-rata harmonik presisi dan recall. Rata-rata harmonik dua bilangan x dan y mendekati bilangan terkecil dari kedua bilangan tersebut. Oleh karena itu, nilai F1-Score yang tinggi memastikan presisi dan recall cukup tinggi.

Berikut cara menghitung F1-Score dari suatu model klasifikasi.

$$F1 = \frac{2(Precision \times Recall)}{Precision + Recall} = \frac{2(TP)}{2 \times TP + FP + FN}$$

6. AUC-ROC

AUC atau Area Under Curve digunakan dalam dugaan dengan Kurva ROC yang merupakan Kurva Karakteristik Operasi Penerima. AUC adalah area di bawah Kurva ROC.



Kurva ROC digambar dengan memplot TPR atau True Positive Rate atau Recall atau Sensitivity (yang kita lihat di atas) pada sumbu y terhadap FPR atau False Positive Rate pada sumbu x . $FPR = 1 - \text{Spesifisitas}$ (yang kita lihat di atas).

Dimana cara menghitung FPR adalah sebagai berikut:

$$FPR = 1 - \frac{TN}{TN + FP} = \frac{FP}{TN + FP}$$

Jika kita menggunakan model acak untuk mengklasifikasikan, ia memiliki probabilitas 50% untuk mengklasifikasikan kelas positif dan negatif dengan benar. Di sini, $AUC = 0,5$. Model sempurna memiliki probabilitas 100% untuk mengklasifikasikan kelas positif dan negatif dengan benar. Di sini, $AUC = 1$. Jadi ketika kita ingin memilih model terbaik, kita menginginkan model yang paling dekat dengan model yang sempurna. Dengan kata lain, model dengan AUC mendekati 1. Ketika kita mengatakan model memiliki skor AUC tinggi, itu berarti kemampuan model untuk memisahkan kelas sangat tinggi (high separability). Ini adalah metrik yang sangat penting yang harus diperiksa saat memilih model klasifikasi.

B. Regression

Berikut merupakan beberapa cara mengukur performa dari suatu model regresi dalam data mining.

1. MAE (Mean Absolute Error)

Ini adalah metrik yang paling sederhana dari semua metrik. Ini diukur dengan mengambil rata-rata perbedaan mutlak antara nilai aktual dan prediksi.

Berikut adalah cara menghitung MAE dalam suatu model regresi.

$$MAE = \frac{1}{n} \sum |y - \hat{y}|$$

2. MSE (Mean Squared Error)

MSE dihitung dengan jumlah kuadrat kesalahan prediksi yang merupakan output nyata dikurangi output yang diprediksi dan kemudian dibagi dengan jumlah titik data. Ini memberi kita angka absolut tentang seberapa jauh hasil prediksi kita menyimpang dari angka sebenarnya. Kita tidak dapat menginterpretasikan banyak pengetahuan dari satu hasil tunggal tetapi memberikan kita bilangan real untuk dibandingkan dengan hasil model lain dan membantu kita memilih model regresi terbaik.

Berikut adalah cara menghitung MSE dalam suatu model regresi.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

3. RMSE (Root Mean Squared Error)

Root Mean Square Error diukur dengan mengambil akar kuadrat dari rata-rata selisih kuadrat antara prediksi dan nilai sebenarnya. Ini mewakili standar deviasi sampel dari perbedaan antara nilai yang diprediksi dan nilai yang diamati (juga disebut residual).

Berikut adalah cara menghitung RMSE dalam suatu model regresi.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{N}}$$