



Smart, Creative and Entrepreneurial

Topik dalam *Image Processing*

Habibullah Akbar, PhD

PERTEMUAN 13

Evaluasi Performa Klasifikasi



Kemampuan Akhir Pertemuan 13

- Setelah mempelajari modul pada sesi ini,
 - Mahasiswa dapat menjelaskan pentingnya melakukan evaluasi performa terhadap model klasifikasi seperti KNN, ANN, dan CNN
 - Mahasiswa dapat menjelaskan pengertian tabel *confusion matrix*.
 - Mahasiswa dapat menjelaskan cara mengukur akurasi, *precision*, *recall* dan *F measure*

Universitas
Esa Unggul

Outline

- Pendahuluan Evaluasi Performa
- *Confusion Matrix*
- Akurasi
- *Precision,*
- *Recall*
- *F measure*

Universitas
Esa Unggul

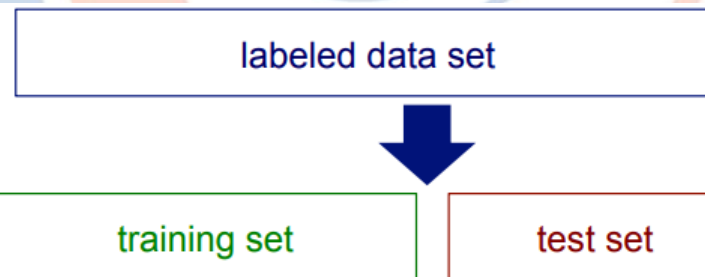
Pendahuluan Evaluasi Performa

- Evaluasi performa pada model klasifikasi seperti KNN, ANN, dan CNN merupakan hal yang penting.
- Misalkan pada dataset penyakit Covid-19, tentunya kita tidak ingin hasil klasifikasi salah, misalnya pasien yang sebenarnya positif Covid-19 justru dideteksi negatif.

Universitas
Esa Unggul

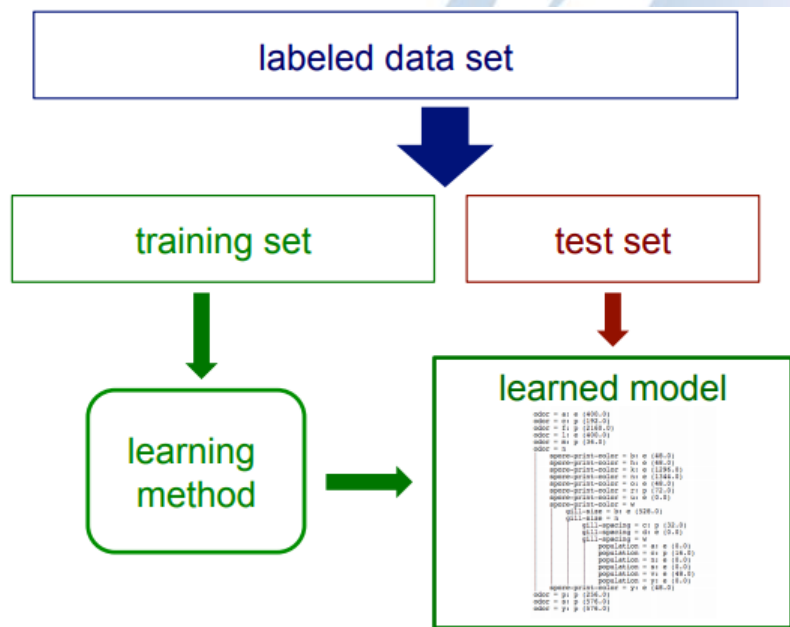
Pendahuluan Evaluasi Performa

- Langkah pertama untuk mengevaluasi adalah kita membagi dataset menjadi *training set* dan *test set*.



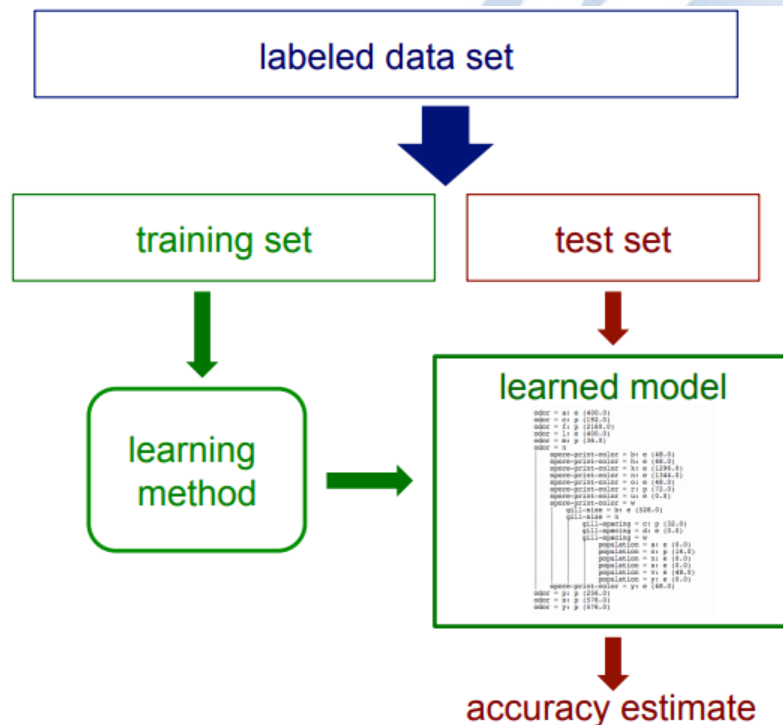
Universitas
Esa Unggul

Pendahuluan Evaluasi Performa



- Berikutnya model klasifikasi (seperti KNN, ANN, dan CNN) akan dilatih menggunakan *training set*.
- Pada contoh VGG (modul 11 dan 12), maka pelatihan bermaksud mencari parameter VGG yang terbaik

Pendahuluan Evaluasi Performa



- Kemudian performa dan model yang dihasilkan perlu dievaluasi agar dapat digunakan dalam bisnis/dunia nyata.

Evaluasi Performa

- Evaluasi dapat dilakukan berdasarkan metrik.
- Contoh metrik adalah akurasi, precision, recall, dan F1 score.
- Metrik-metrik tersebut dapat diturunkan berdasarkan *confusion matrix*.

Universitas
Esa Unggul

Confusion Matrix

- Misalkan kita ingin mendeteksi penyakit Covid-19.
- Matrik confusion adalah tabel untuk mengevaluasi performa suatu metode klasifikasi.

		HASIL PREDIKSI	
		positif	negatif
LABEL SEBENARNYA	positif	true positives (TP)	false negatives (FN)
	negatif	false positives (FP)	true negatives (TN)

Confusion Matrix

- Label positif menunjukkan pasien sebenarnya menderita penyakit Covid-19. sedangkan label negatif menunjukkan pasien sebenarnya tidak terkena Covid-19. Pada tabel disebut LABEL YANG SEHARUSNYA.
- HASIL PREDIKSI dari metode klasifikasi yang di-training dapat saja salah.

		HASIL PREDIKSI	
		positif	negatif
LABEL YANG SEHARUSNYA	positif	true positives (TP)	false negatives (FN)
	negatif	false positives (FP)	true negatives (TN)

Confusion Matrix

- Ketika hasil prediksi metode klasifikasi mengatakan positif dan label yang seharusnya adalah positif, maka kita matrix ini menyatakan *true positif* (memang benar positif). Demikian juga, jika hasil prediksi negatif dan memang label negatif, kita katakan *true negatif* (sesuai)
- Namun jika hasil prediksi menyatakan pasien positif Covid-19 padahal sebenarnya negatif (labelnya) maka kejadian ini disebut *false positif* (yaitu kesalahan dalam memprediksi positif) dan jika sebaliknya maka disebut *false negatif*.

		HASIL PREDIKSI	
		positif	negatif
LABEL SEHARUSNYA	positif	true positives (TP)	false negatives (FN)
	negatif	false positives (FP)	true negatives (TN)

Contoh

- Pada contoh tabel, terjadi kesalahan prediksi sebanyak 10 data. Empat pasien yang seharusnya negative (labelnya) justru diprediksi positif Covid-19. Akibatnya, pasien bisa dikarantina padahal seharusnya tidak perlu.
- Sebaliknya, 6 pasien yang memang terkena Covid-19 diprediksi negatif. Akibatnya, penyebaran bisa semakin luas karena pasien dapat berkeliaran bebas.

		HASIL PREDIKSI	
		positif	negatif
LABEL SEHARUSNYA	positif	60	6
	negatif	4	30

Metrik

- Dari *confusion matrix*, dapat diturunkan beberapa metrik untuk mengukur performa model klasifikasi seperti:
 - Akurasi
 - Precision,
 - Recall
 - F measure

Universitas
Esa Unggul

Akurasi

- Akurasi adalah nilai total dari prediksi yang tepat yaitu dari *True Positive* (TP) dan *True Negative* (TN). Perhitungannya adalah sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- Kelemahan akurasi adalah ketika dataset tidak seimbang. Misal, dataset mengandung data berlabel positif terlalu banyak dan hanya sedikit data berlabel negatif. Untuk mengatasi, dapat digunakan *F measure*.

Precision

- Presisi adalah perbandingan dari prediksi positif yang tepat terhadap total prediksi yang positif. Perhitungannya adalah sebagai berikut:

$$Precision = \frac{TP}{TP + FP}$$

- Presisi dapat diartikan sebagai ukuran tentang berapa banyak pasien Covid-19 yang diprediksi positif berdasarkan label yang memang seharusnya positif.

Recall

- *Recall* (disebut juga sensitifitas) adalah perbandingan dari prediksi positif yang tepat terhadap label yang memang seharusnya positif .

$$\text{Recall/Sensitivity} = \frac{TP}{TP + FN}$$

- *Recall* digunakan untuk mengevaluasi resiko pasien yang sebenarnya positif Covid-19 justru dideteksi negatif.

Esa Unggul

F measure

- Ukuran (*measure*) F adalah rata-rata terhadap precision dan recall yang dibobotkan. Ukuran F lebih tepat digunakan ketika dataset yang dilabelkan tidak seimbang.

$$F - Measure = \frac{2 * Precision * Recall}{Precision + Recall}$$

Universitas
Esa Unggul

Catatan

- Selain membagi dataset menjadi 2 yaitu training set dan test set, terdapat metode evaluasi performa lainnya seperti:
 - Cross validation
 - Kurva *receiver operating characteristic*

Universitas
Esa Unggul

Referensi

1. GONZALEZ, R. C., AND WOODS, R. E. 2018. Digital image processing, 4th Global Edition, Pearson, New York
2. SOLOMON, C. AND BRECKON, T. 2011. Fundamentals of Digital Image Processing: A practical approach with examples in Matlab. John Wiley & Sons.
3. Tan, Steinbach, and Kumar, Introduction to Data Mining
4. <http://www-users.cs.umn.edu/~kumar/dmbook/index.php>
5. Han, Kamber, and Pei, Data mining: Concepts and Techniques
http://hanj.cs.illinois.edu/bk3/bk3_slidesindex.htm
6. Syarif, I., Prugel-Bennett, A., & Wills, G. (2016). SVM parameter optimization using grid search and genetic algorithm to improve classification performance. Telkomnika, 14(4), 1502.