



Computer Science Department
Bogor Agricultural University
<http://cs.ipb.ac.id/>

Pengenalan Pola

Kuliah 14

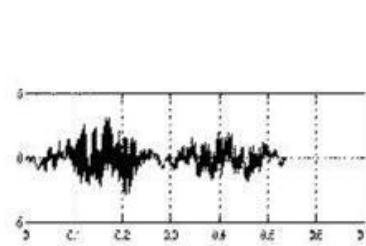
Outline

- Pengenalan Pola
- Sistem Pengenalan Pola



Pengenalan Pola

- Pola (Objek)
 - Entitas yang terdefinisi dan dapat diidentifikasi melalui ciri-cirinya (*features*)
- Ciri-ciri tsb membedakan suatu pola dengan pola yang lain



Pengenalan Pola

- Ciri yang baik adalah ciri yang memiliki daya pembeda yang tinggi
- Ciri pada suatu pola diperoleh dari hasil pengukuran terhadap objek

Pola	Ciri
huruf	tinggi, tebal, titik sudut, lengkungan garis, dll
suara	amplitudo, frekuensi, nada, intonasi, warna, dll
tanda tangan	panjang, kerumitan, tekanan, dll
sidik jari	lengkungan, jumlah garis, dll



Pengenalan Pola

- Pola yg terdapat di dalam citra digital, ciri-cirinya berasal dari informasi:
 - Spasial : intensitas *pixel*, histogram, ketetanggaan, ...
 - Tepi : arah, kekuatan, ...
 - Kontur : garis, elips, lingkaran
 - Wilayah/bentuk : keliling, luas, ...
 - Hasil transformasi Fourier : frekuensi, ...

Kelas Pola

- Kelas Pola (Kategori) :
 - Himpunan pola yang memiliki atribut tertentu
- Contoh :
 - Pada pola huruf, terdapat beberapa jenis huruf, yakni huruf a, b, c, d, dst.
 - Huruf a, b, c, d merupakan kelas pola dari pola huruf.

Pengenalan Pola vs Persepsi Manusia

- Manusia telah dianugerahi kemampuan untuk menerima rangsangan (indera) dari lingkungan dan memberikan aksi terhadap apa yang diamati
 - Mengenali wajah
 - Memahami kata yang diucapkan
 - Membaca tulisan tangan
 - Membedakan makanan segar dari baunya

Pengenalan Pola vs Persepsi Manusia

- Pengenalan pola bertujuan :
 - menentukan kelompok atau kategori pola berdasarkan ciri-ciri yang dimiliki oleh pola tersebut
- Tantangan:
 - Menjadikan mesin (komputer) memiliki kemampuan yang mirip dengan manusia



Contoh Aplikasi Pengenalan Pola

Problem Domain	Application	Input Pattern	Pattern Classes
Document image analysis	Optical character recognition	Document image	Characters, words
Document classification	Internet search	Text document	Semantic categories
Document classification	Junk mail filtering	Email	Junk/non-junk
Multimedia database retrieval	Internet search	Video clip	Video genres
Speech recognition	Telephone directory assistance	Speech waveform	Spoken words
Natural language processing	Information extraction	Sentences	Parts of speech
Biometric recognition	Personal identification	Face, iris, fingerprint	Authorized users for access control
Medical	Computer aided diagnosis	Microscopic image	Cancerous/healthy cell
Military	Automatic target recognition	Optical or infrared image	Target type
Industrial automation	Printed circuit board inspection	Intensity or range image	Defective/non-defective product
Industrial automation	Fruit sorting	Images taken on a conveyor belt	Grade of quality
Remote sensing	Forecasting crop yield	Multispectral image	Land use categories
Bioinformatics	Sequence analysis	DNA sequence	Known types of genes
Data mining	Searching for meaningful patterns	Points in multidimensional space	Compact and well-separated clusters

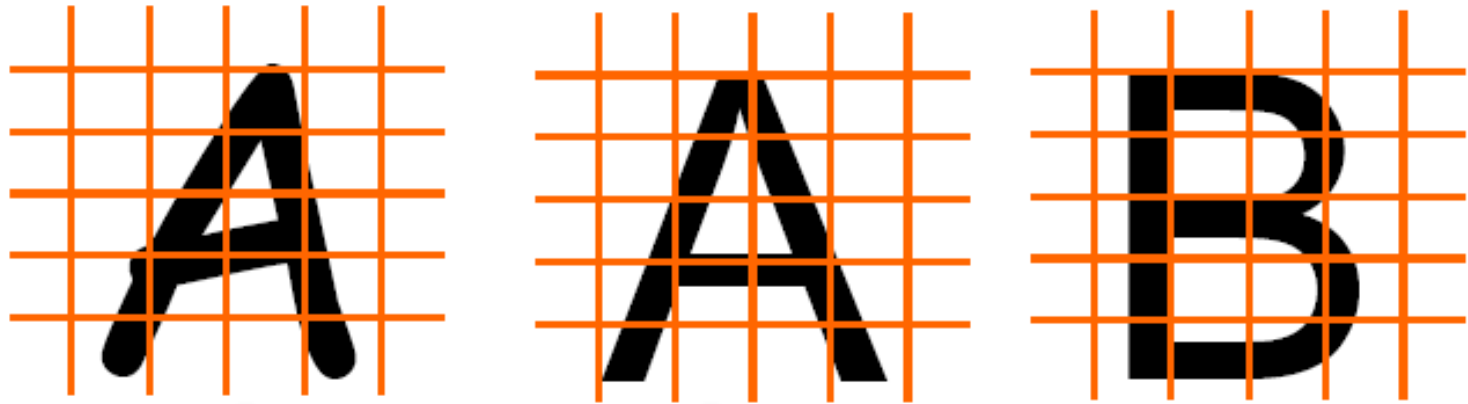


Pendekatan Pengenalan Pola

- Statistical :
 - Berdasarkan model statistik dari pola dan kelas pola yang diberikan
- Structural :
 - Kelas pola direpresentasikan oleh struktur formal seperti grammer, string, automata, dll.

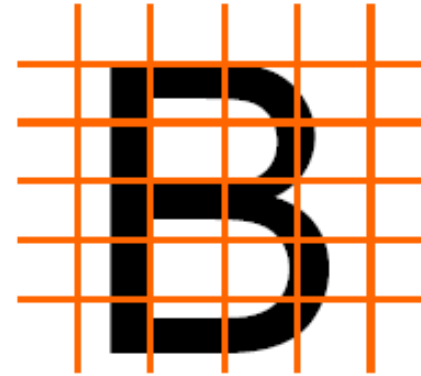
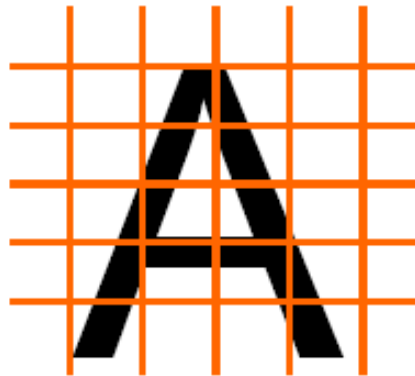
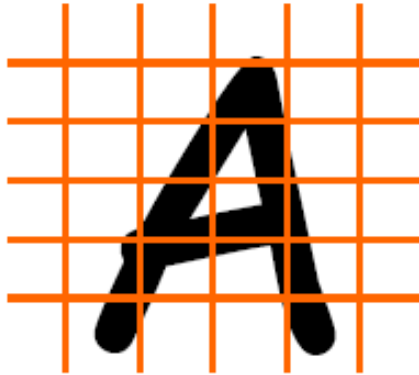


Pendekatan Statistik



**Membandingkan
Grid per Grid**

Pendekatan Statistik

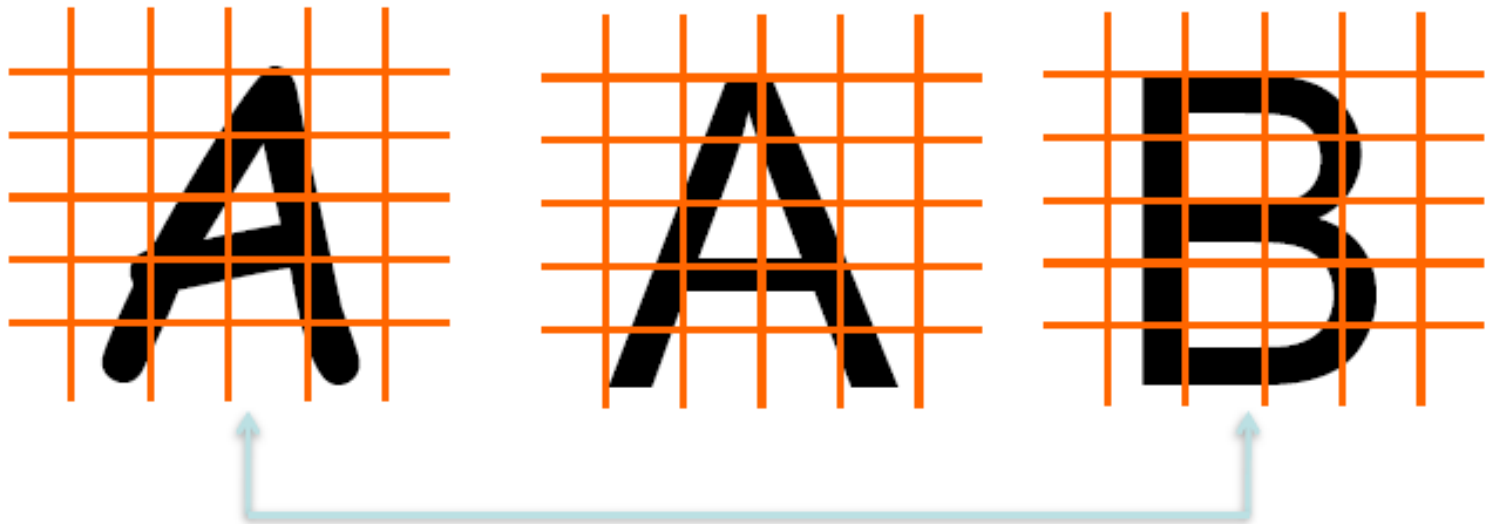


0	0	1	0
0	0	1	0
0	1	1	1
1	0	0	1
1	0	0	1

0	1	1	0
0	1	1	0
0	1	1	0
1	0	0	1
1	0	0	1

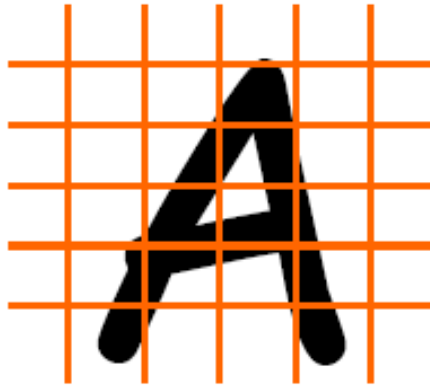
Jumlah grid
yang tidak
sesuai = 3

Pendekatan Statistik

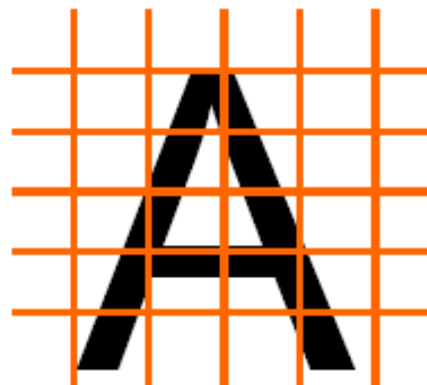


**Membandingkan
Grid per Grid**

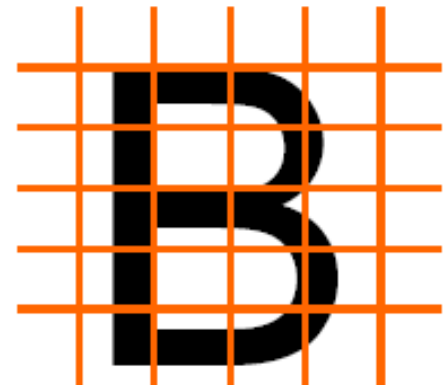
Pendekatan Statistik



0	0	1	0
0	0	1	0
0	1	1	1
1	0	0	1
1	0	0	1



Jumlah grid
yang tidak
sesuai = 10



1	1	1	0
0	1	0	1
0	1	1	1
0	1	0	1
1	1	1	0

Pendekatan Statistik

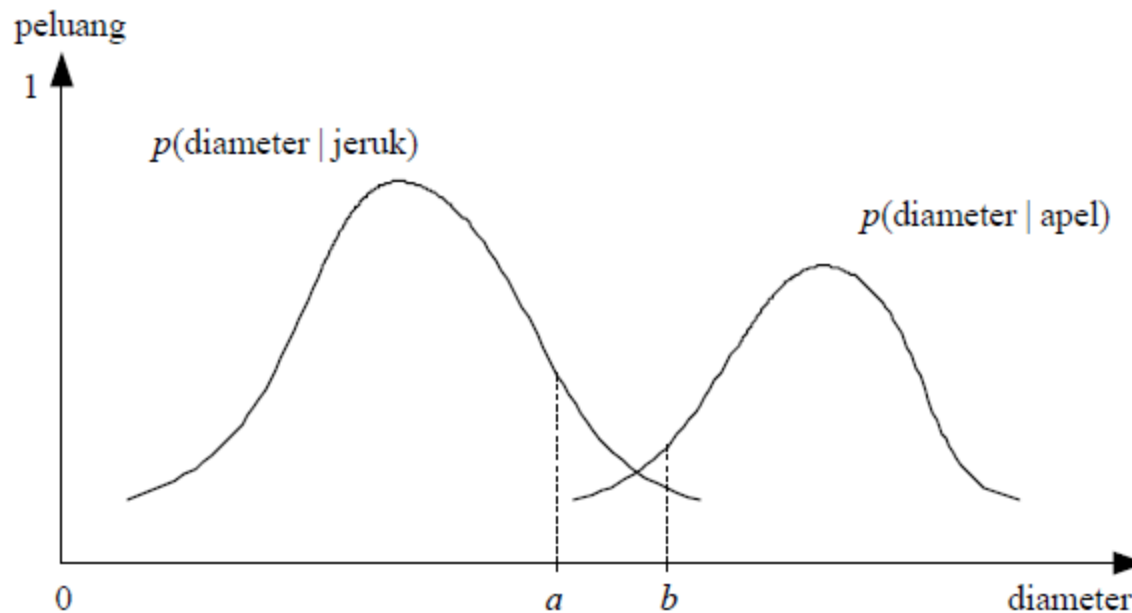
- Ciri-ciri yang dimiliki oleh suatu pola ditentukan distribusi statistiknya.
- Pola yang berbeda memiliki distribusi yang berbeda pula.
- Dengan menggunakan teori keputusan di dalam statistik, kita menggunakan distribusi ciri untuk mengklasifikasikan pola.

Pendekatan Statistik

- Misalkan ada N pola yang dikenali, yaitu $w_1, w_2, w_3, \dots, w_N$ dan fungsi peluang atau kerapatan dari ciri-ciri pada pola diketahui.
- Jika x merupakan hasil pengukuran ciri-ciri, maka $p(x|w_i)$, $i = 1, 2, \dots, N$ dapat dihitung.

Pendekatan Statistik

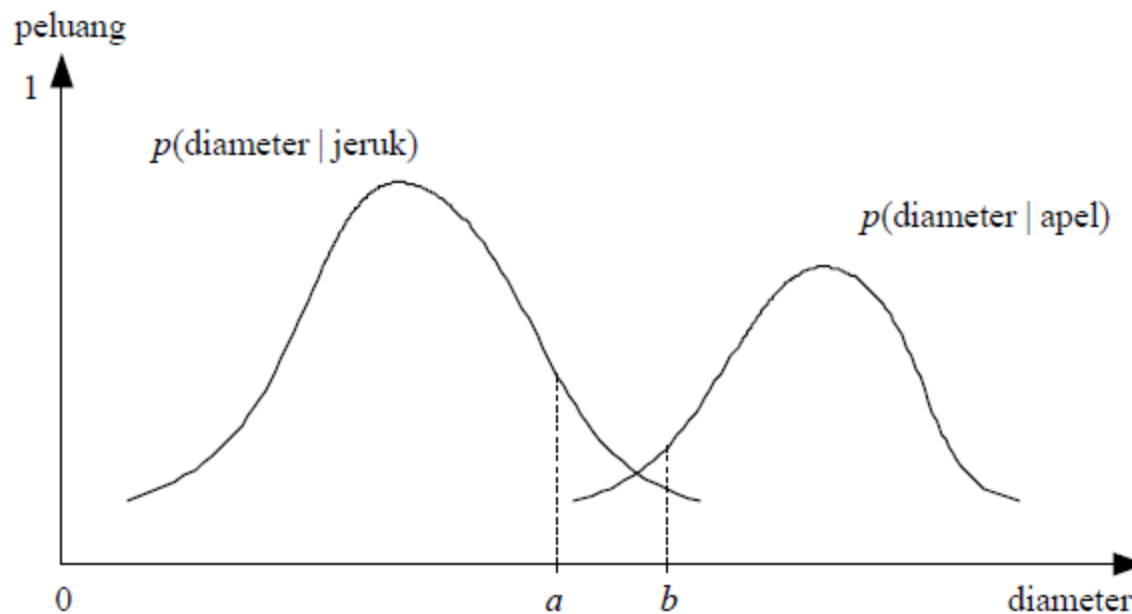
- Sebagai contoh, misalkan diketahui fungsi kerapatan dari diameter buah jeruk dan apel



Gambar 15.1. Grafik fungsi kerapatan dari ciri diameter jeruk dan apel.

Pendekatan Statistik

- Jika sebuah objek diukur dan diperoleh diameternya adalah a cm, maka kita mengklasifikasikan objek tersebut sebagai "jeruk", karena $p(a|\text{jeruk}) > p(a|\text{apel})$



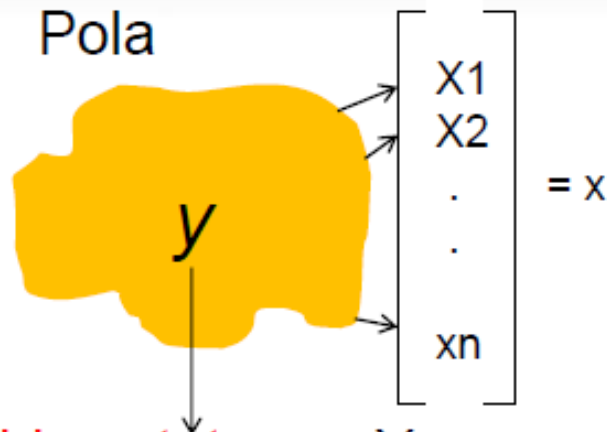
Gambar 15.1. Grafik fungsi kerapatan dari ciri diameter jeruk dan apel.

Pengenalan Pola

- Dua Tahap
 - **Learning (training process)**
 - beberapa contoh citra dipelajari untuk menentukan ciri yang akan digunakan dalam proses pengenalan serta prosedur klasifikasinya.
 - **Detection (testing process)**
 - citra diambil cirinya kemudian ditentukan kelas kelompoknya
- Waktu Learning lebih besar
- Sulit untuk belajar, tetapi sekali terpelajar sistem akan menjadi "**natural**"
- Dapat menggunakan metode AI :
 - Neural Network
 - Machine Learning



Konsep Dasar



Feature vector $x \in X$

- Vector dari hasil pengamatan (pengukuran).
- x adalah sebuah titik dalam ruang vektor X

Hidden state $y \in Y$

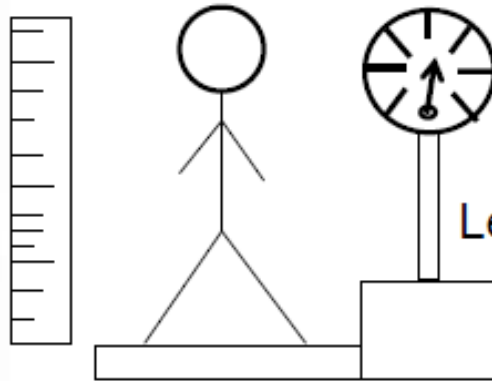
- Tidak dapat diukur secara langsung.
- Pola dengan hidden state yang sama terdapat dalam kelas yang sama

Task

- Mendesain classifier (decision rule) $q : X \rightarrow Y$ yang menentukan hidden state berdasarkan pengamatan

Contoh

Tinggi



$\begin{bmatrix} X1 \\ X2 \end{bmatrix}$

$= x$

Task: Human recognition

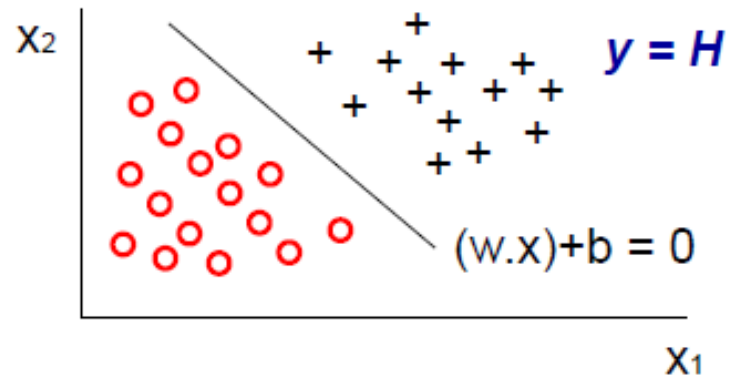
Himpunan hidden state $Y = \{H, J\}$

Ruang vektor $X = \mathbb{R}^2$

Training examples : $\{(x_1, y_1), \dots, (x_i, y_i)\}$

Linier classifier :

$$q(x) = \begin{cases} H & \text{if } (w \cdot x) + b \geq 0 \\ J & \text{if } (w \cdot x) + b < 0 \end{cases}$$



X_1

Learning

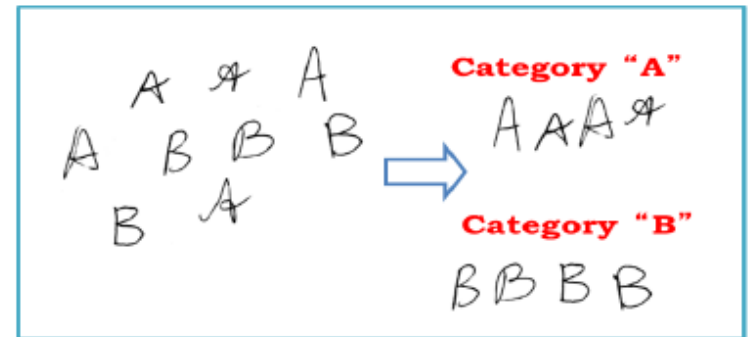
- Bagaimana mesin dapat belajar aturan dari data.
 - **Supervised learning :**
 - User menyediakan label kategori untuk masing-masing pola dalam data training.
 - **Unsupervised learning :**
 - Sistem membentuk cluster atau pengelompokan secara alami dari pola yang dimasukan.



Learning

- **Classification**

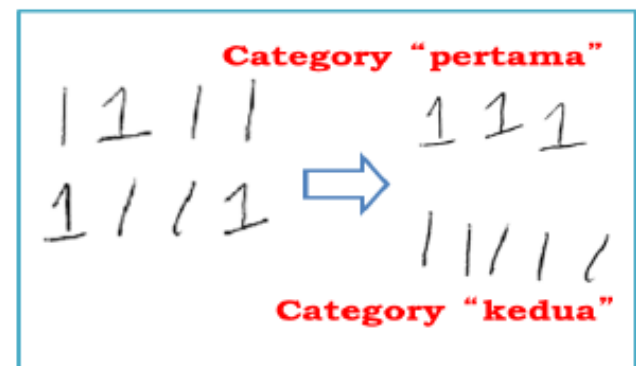
- Kategori/label kelasnya diketahui



Classification
(Supervised Classification)

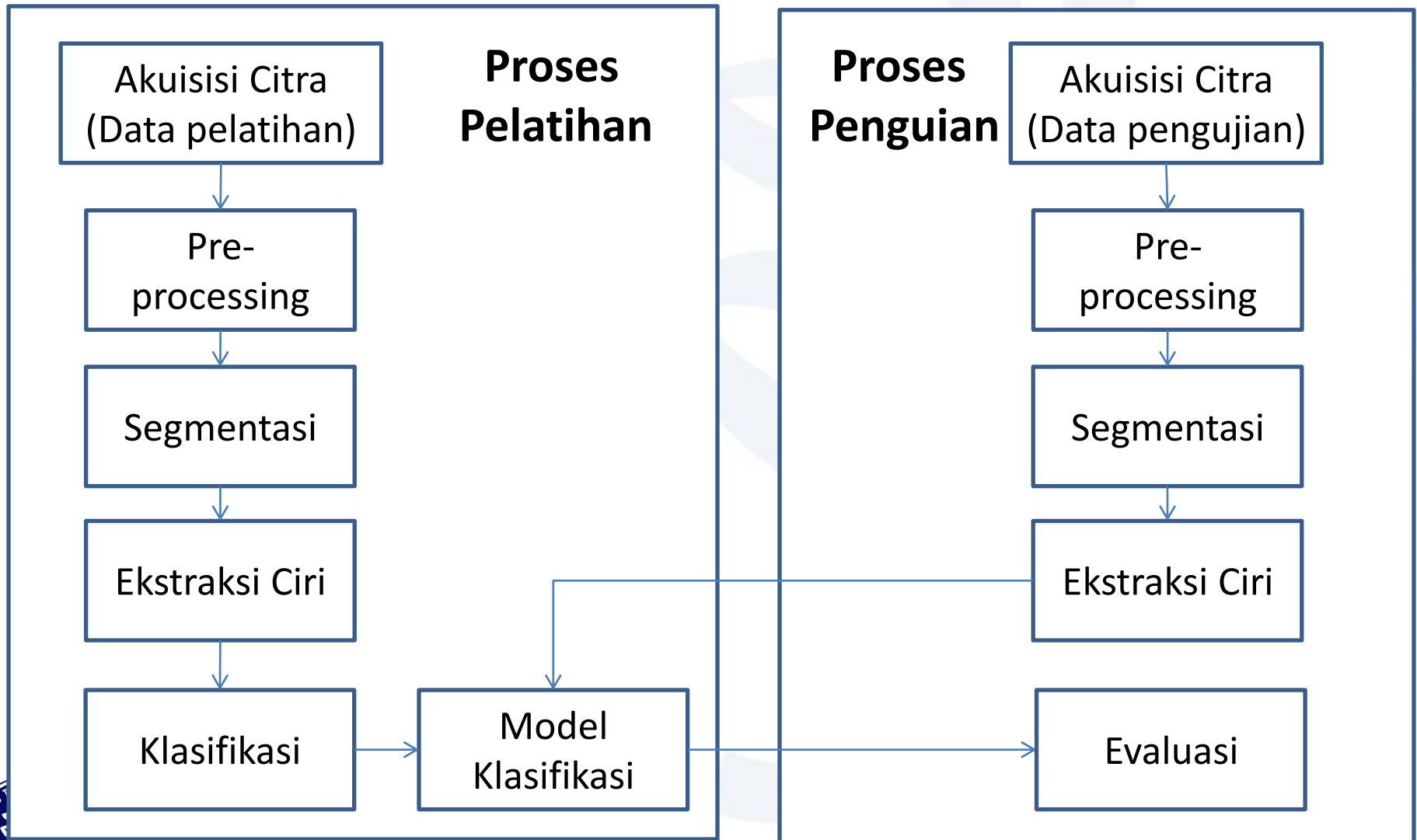
- **Clustering**

- Kategori/label kelompoknya tidak diketahui dan biasanya dinamai dengan kelas pertama, kedua dst.



Clustering
(Unsupervised Classification)

Proses Pengenalan Pola (klasifikasi)



Proses Pengenalan Pola (klasifikasi)

- **Akuisisi Citra**

- Proses untuk mendapatkan citra digital
- Alat :
 - kamera digital, scanner, mikroskop digital, webcam, dll.
- Yang perlu diperhatikan :
 - Resolusi, jarak dan sudut pengambilan, pencahayaan, perbesaran, format citra hasil akuisisi

Proses Pengenalan Pola (klasifikasi)

- **Pre-processing**

- Perbaikan kualitas citra (*image enhancement, image restoration*)
- Metode perbaikan kualitas citra di antaranya adalah:
 - *contrast stretching,*
 - *filtering (median filter, low pass filter, high pass filter, dsb).*

Proses Pengenalan Pola (klasifikasi)

- **Segmentasi**

- Proses pemisahan objek yang akan diamati (*foreground*) dengan objek yang tidak akan diamati (*background*) dalam suatu gambar
- Berupa citra biner :
 - foreground (bernilai 1), background (bernilai 0)
- Metode segmentasi citra di antaranya :
 - Thresholding
 - Multi-thresholding
 - Deteksi tepi
 - Active countour
 - K-means clustering
 - Fuzzy k-means clustering
 - dll



Proses Pengenalan Pola (klasifikasi)

- **Ekstraksi Ciri**

- tahapan mengekstrak ciri dari pola/objek di dalam citra yang ingin dikenali/dibedakan dengan pola/objek lainnya.

- Ciri-ciri umum yang diekstrak :

- Ekstraksi ciri bentuk :
 - eccentricity, convex hull, hough transform, dll.
- Ekstraksi ciri ukuran dan geometri:
 - luas, keliling, panjang, lebar, diameter, dll.
- Ekstraksi ciri tekstur :
 - Local Binary Pattern, Gray Level Co-occurrence Matrix, Gabor Filter, dll.
- Ekstraksi ciri warna :
 - Color Histogram, Fuzzy color histogram, color moment, dll



Proses Pengenalan Pola (klasifikasi)

- **Klasifikasi**

- Menggunakan fitur untuk mendapatkan model pembelajaran untuk menetapkan pola pada suatu kelas/kategori tertentu

- Metode klasifikasi yang biasa digunakan:

- Neural Networks (Jaringan saraf tiruan)
 - Decision tree (pohon keputusan)
 - Support Vector Machine (SVM)
 - K-Nearest Neighbour (KNN)
 - Naive Bayes



Proses Pengenalan Pola (klasifikasi)

- **Evaluasi**

- Ukuran tingkat keberhasilan klasifikasi/model klasifikasi yang dibentuk terhadap citra uji
- Salah satu ukuran evaluasi pada proses klasifikasi, yakni akurasi.
- Kita dapat menggunakan **Confusion Matrix** untuk menghitung akurasi.

Proses Pengenalan Pola (klasifikasi)

- Confusion matrix

		predicted condition	
		prediction positive	prediction negative
true condition	condition positive	True Positive (TP)	False Negative (FN) (type II error)
	condition negative	False Positive (FP) (Type I error)	True Negative (TN)

$$\text{Accuracy} = \frac{\Sigma \text{TP} + \Sigma \text{TN}}{\Sigma \text{total population}}$$

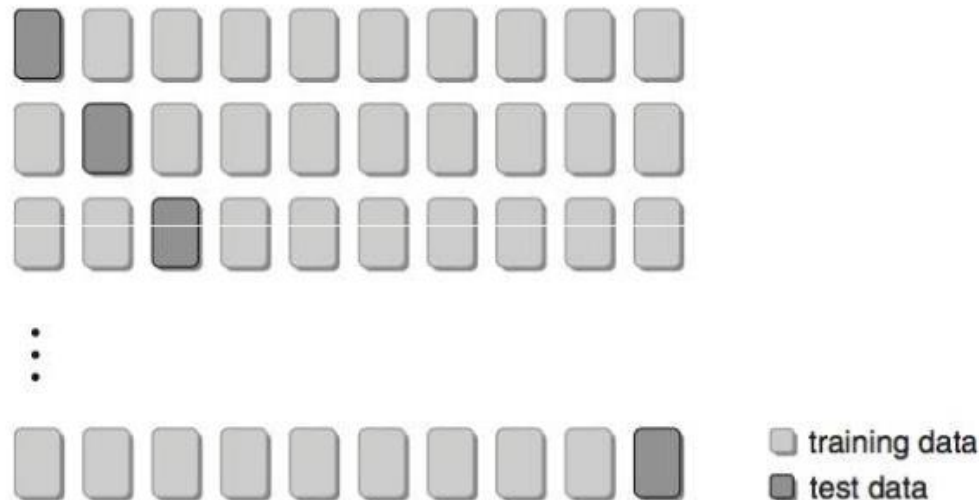
Proses Pengenalan Pola (klasifikasi)

- Untuk dapat memvalidasi keakuratan sebuah model perlu dilakukan tahapan ***cross validation***.
- Salah satu metode cross-validation yang populer : **K-Fold Cross Validation**
- Dalam teknik ini dataset dibagi menjadi sejumlah K-buah partisi secara acak.



Proses Pengenalan Pola (klasifikasi)

- Kemudian dilakukan sejumlah K-kali eksperimen, dimana masing-masing eksperimen menggunakan data partisi ke-K sebagai data testing dan memanfaatkan sisa partisi lainnya sebagai data training.



Contoh : 5 fold cross validation

Dataset

K1	K2	K3	K4	K5
----	----	----	----	----

Data Eksperimen

Eksperimen ke	Data Training	Data Testing
1	K2,K3,K4,K5	K1
2	K1,K3,K4,K5	K2
3	K1,K2,K4,K5	K3
4	K1,K2,K3,K5	K4
5	K1,K2,K3,K4	K5