

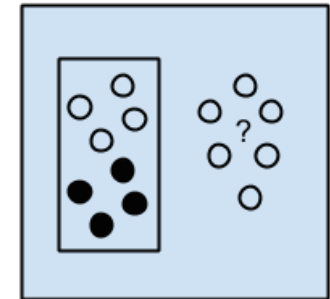
Pertemuan 4

Supervised Learning Klasifikasi dan Regresi

Supervised vs Unsupervised

Supervised

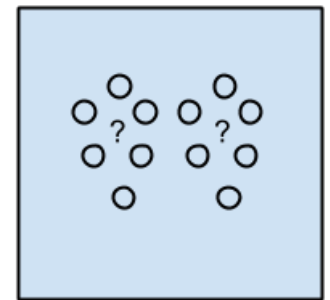
- Memiliki target
- Temukan fungsi yang dapat memetakan data pada targetnya
- Menemukan pola yang menghubungkan atribut dengan targetnya



Supervised Learning Algorithms

Unsupervised

- Tidak memiliki target
- Menemukan struktur data yang mendasarinya
- Tidak memprediksikan secara spesifik, hanya mengelompokkan saja



Unsupervised Learning Algorithms

Klasifikasi

- Klasifikasi adalah menentukan sebuah record data baru ke salah satu dari beberapa kategori (atau klas) yang telah didefinisikan sebelumnya.
- Termasuk kedalam “*supervised learning*”.

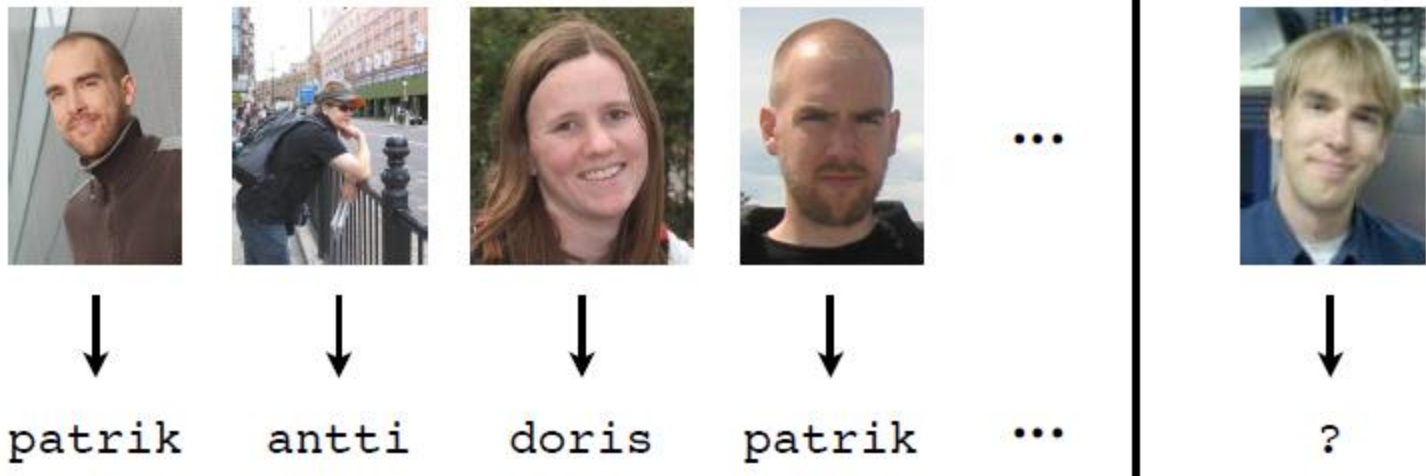
Contoh Klasifikasi Spam Filter

- X adalah kumpulan email
- Y adalah target/kelas apakah { spam atau non-spam }

From: medshop@spam.com Subject: viagra cheap meds...	spam
From: my.professor@helsinki.fi Subject: important information here's how to ace the test...	non-spam
⋮	⋮
From: mike@example.org Subject: you need to see this how to win \$1,000,000...	?

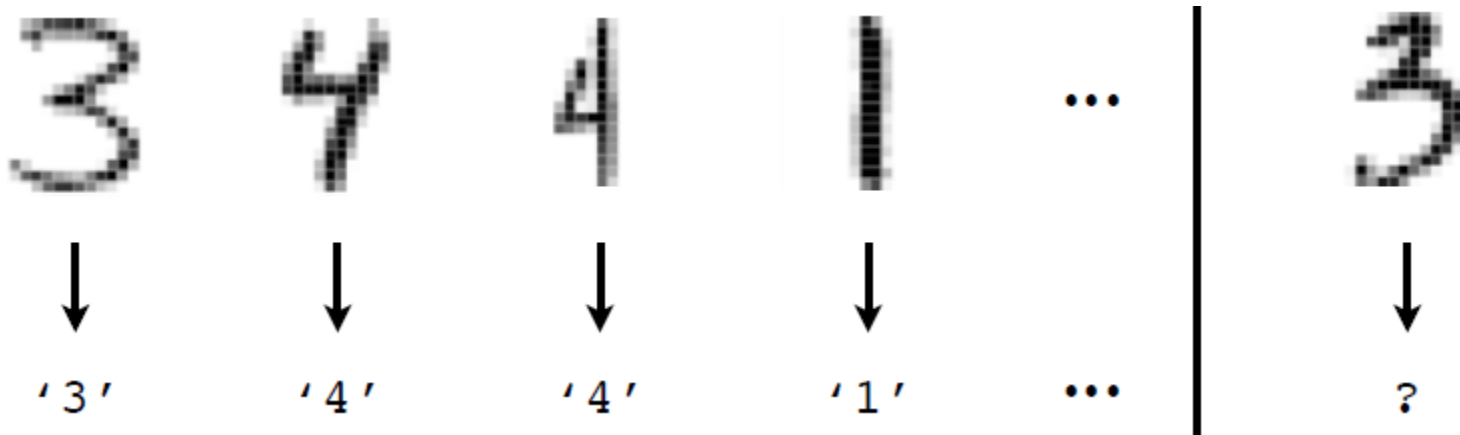
Contoh Klasifikasi Face Recognition

- X adalah kumpulan gambar foto wajah
- Y adalah target/kelas



Contoh Klasifikasi Pengenalan Tulisan Tangan

- X adalah kumpulan gambar tulisan tangan berisi angka
- Y adalah target/kelas berupa angka $\{ 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 \}$



Contoh Klasifikasi

Klasifikasi Kelulusan Mahasiswa

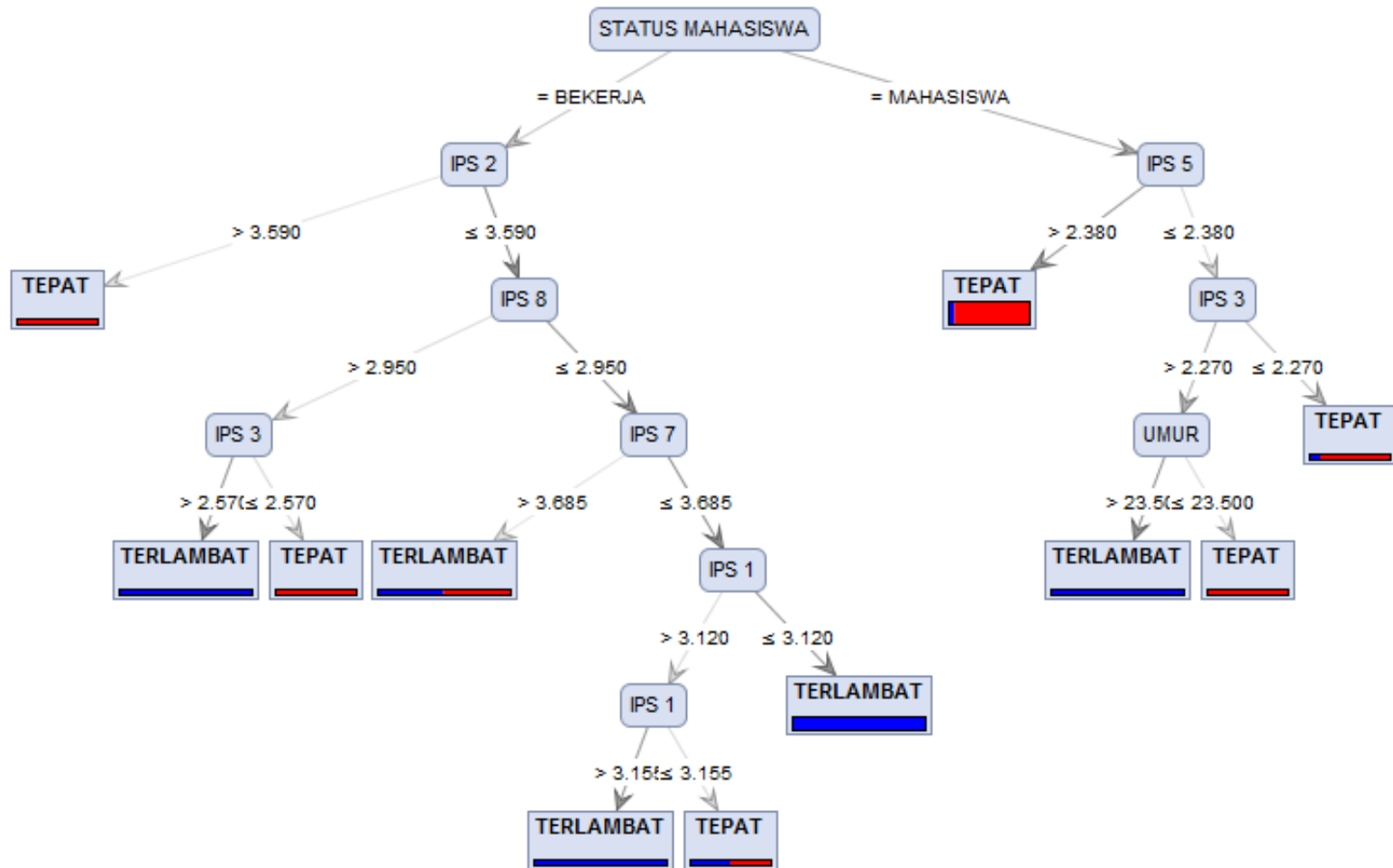
Label
↓

NIM	Gender	Nilai UN	Asal Sekolah	IPS1	IPS2	IPS3	IPS 4	...	Lulus Tepat Waktu
10001	L	28	SMAN 2	3.3	3.6	2.89	2.9		Ya
10002	P	27	SMA DK	4.0	3.2	3.8	3.7		Tidak
10003	P	24	SMAN 1	2.7	3.4	4.0	3.5		Tidak
10004	L	26.4	SMAN 3	3.2	2.7	3.6	3.4		Ya
...									
...									
11000	L	23.4	SMAN 5	3.3	2.8	3.1	3.2		Ya

Pembelajaran dengan
Metode Klasifikasi (C4.5)
↓

Klasifikasi Kelulusan Mahasiswa

Hasil Model Klasifikasi Menggunakan Decision Tree



Regresi

- Memprediksi nilai dari suatu variabel kontinyu yang diberikan berdasarkan nilai dari variabel yang lain, dengan mengasumsikan sebuah model ketergantungan linier atau nonlinier
- Termasuk kedalam "supervised learning"

Contoh Regresi

- Memprediksi jumlah penjualan produk baru berdasarkan pada belanja promosi/iklan
- Memprediksi kecepatan angin sebagai suatu fungsi suhu, kelembaban, tekanan udara, dsb.
- *Time series prediction dari indeks stock market*

Simulasi Regresi (Estimasi)

Customer	Jumlah Pesanan (P)	Jumlah Traffic Light (T)	Jarak (J)	Waktu Tempuh (T)
1	3	3	3	16
2	1	7	4	20
3	2	4	6	18
4	4	6	8	36
...				
1000	2	4	2	12

Label

Pembelajaran dengan
Metode Estimasi (*Regresi Linier*)

$$\text{Waktu Tempuh (T)} = 0.48P + 0.23T + 0.5J$$

Pengetahuan

Simulasi Regresi (Prediksi)

Label



Row No.	Close	Date	Open	High	Low	Volume
1	1286.570	Apr 11, 2006	1296.600	1300.710	1282.960	2232880000
2	1288.120	Apr 12, 2006	1286.570	1290.930	1286.450	1938100000
3	1289.120	Apr 13, 2006	1288.120	1292.090	1283.370	1891940000
4	1285.330	Apr 17, 2006	1289.120	1292.450	1280.740	1794650000
5	1307.280	Apr 18, 2006	1285.330	1309.020	1285.330	2595440000
6	1309.930	Apr 19, 2006	1307.650	1310.390	1302.790	2447310000
7	1311.460	Apr 20, 2006	1309.930	1318.160	1306.380	2512920000
8	1311.280	Apr 21, 2006	1311.460	1317.670	1306.590	2392630000
9	1308.110	Apr 24, 2006	1311.280	1311.280	1303.790	2117330000
10	1301.740	Apr 25, 2006	1308.110	1310.790	1299.170	2366380000
11	1305.410	Apr 26, 2006	1301.740	1310.970	1301.740	2502690000
12	1309.720	Apr 27, 2006	1305.410	1315	1295.570	2772010000
13	1310.610	Apr 28, 2006	1309.720	1316.040	1306.160	2419920000

Dataset harga saham dalam bentuk **time series** (rentet waktu)



Pembelajaran dengan
Metode Prediksi (*Neural Network*)