

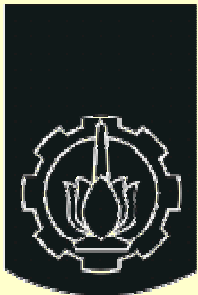
Studi kasus aplikasi Data Mining

Deteksi mahasiswa yang berbakat dan yang bermasalah

Ali Ridho Barakbah

Soft **C**omputing Research Group
Electronic Engineering Polytechnic Institute of Surabaya
(**SCOPE**)

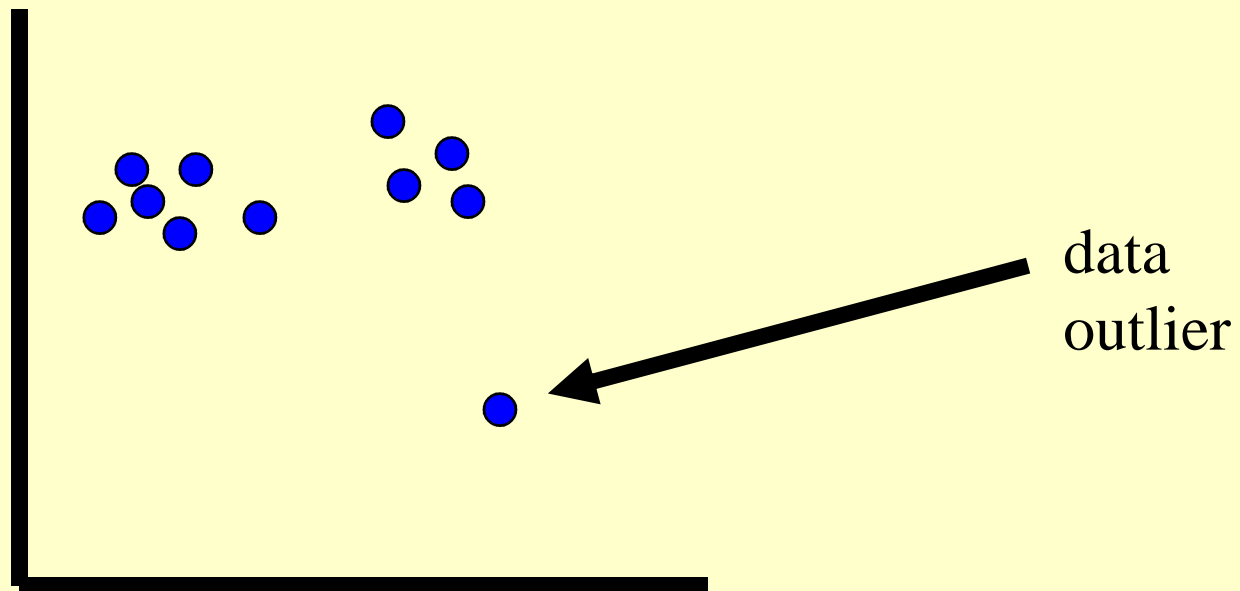
Institut Teknologi Sepuluh Nopember Surabaya



Outlier data

- Adalah data memisahkan diri (berada jauh) dari sebagian besar data yang lain
- Kata lain dari outlier → anomaly, noisy
- Data anomaly biasanya bermanfaat untuk mengevaluasi sistem
- Data noisy perlu dihaluskan atau dibuang

Ilustrasi



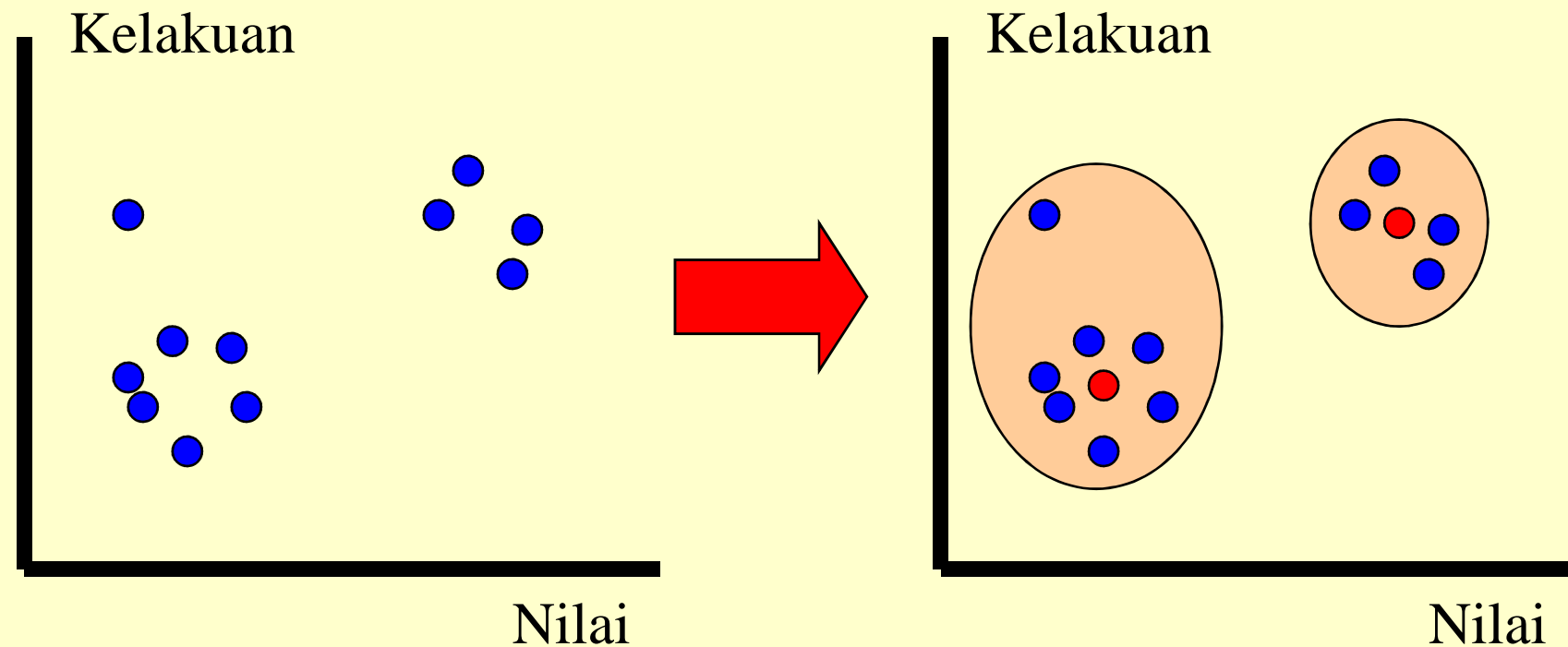
Case description

- Atribut : nilai mata kuliah, kelakuan, keaktifan
- Data : daftar mahasiswa
- Jumlah cluster = 2 (baik dan tidak)

Contoh meta data

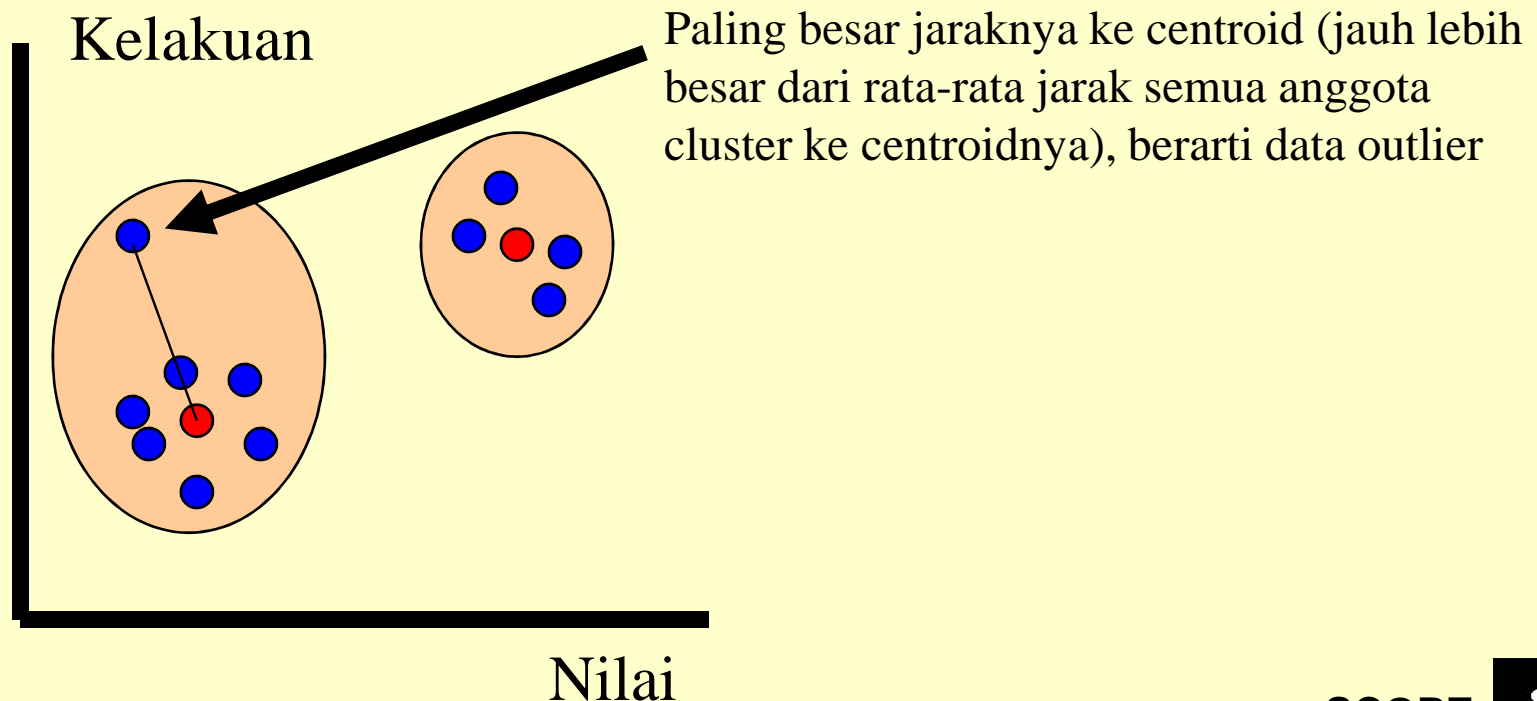
	Nilai OOP	Nilai Jaringan	Nilai AI	Kelakuan	Keaktifan
Agus					
Wati					
Andi					
Tuti					

Clustering



Deteksi outlier

- Hitung jarak setiap anggota cluster ke centroidnya.
- Hitung rata-rata jarak diatas
- Jika selisih jarak suatu data terlalu besar (berikan threshold) terhadap rata-rata jarak, maka data itu adalah outlier



Kasus 1

- Data 1 → {80 60 75 73 81 66}
- Data 2 → {70 84 88 90 65 60}
- Data 3 → {60 40 55 58 47 49}
- Data 4 → {60 65 60 70 68 72}
- Data 5 → {60 35 40 20 56 57}
- Data 6 → {95 98 85 87 93 95}
- Data 7 → {55 70 53 64 74 77}
- Data 8 → {50 53 57 63 58 40}
- Data 9 → {62 64 53 58 40 45}

Carilah data outlier

Kasus 2

- Carilah data outlier pada nilai mahasiswa IT kelas 4D4A
- Mata kuliah yang dilibatkan: OOP, RPL, Database 1, Jaringan dan AI.