



# Data Mining

---

ARIK SOFAN TOHIR, M.KOM

# Mengapa Data Mining?

## Tinjauan Bisnis

---

- ❖ Pengumpulan informasi dari penyimpanan data besar
  - Penggunaan data POS (bar code scanner, RFID, smart card, and smart tech) untuk keperluan :
    - EOS (Electronic Ordering System)
    - Pemilihan item produk berdasarkan karakteristik
    - Penataan tata letak produk

# Mengapa Data Mining?

## Tinjauan Bisnis

---

Segmentasi  
Pelanggan

- Data transaksi
- Rute yang harus dilalui

Sistem  
Rekomendasi  
Produk

- Catatan layanan pelanggan

# Mengapa Data Mining?

## Tinjauan Bisnis

---

Proses layanan  
atau peningkatan  
kualitas Pelanggan

- Transaksi Bank/Kartu Kredit
- Bank/Kartu Kredit transaksi

Pelanggan  
potensial

- Penentuan pemberian kredit
- Penentuan wilayah pemasaran

# Mengapa Data Mining? Tinjauan Bisnis

---

- Teknik Data Mining dapat diaplikasikan secara luas pada segala jenis bidang
  - Customer Profiling
  - Perencanaan target pasar
  - Layout toko
  - Fraud detection (deteksi penipuan)

# Mengapa Data Mining?

## Tinjauan Bisnis

---

- ❑ Menyediakan data kepada manajemen untuk mengambil keputusan strategis dan bisnis lebih baik
- ❑ Mayoritas aplikasi melibatkan serangkaian aktivitas
  - ❑ Query Report
  - ❑ OLAP
  - ❑ DSS
  - ❑ Forecasting
  - ❑ Analisi statistik



# Mengapa Data Mining?

## Tinjauan Keilmuan

---

- Pengumpulan dan penyimpanan data dalam kecepatan tinggi, baik melalui sensor, telescope, gen, maupun simulasi scientific



# Mengapa Data Mining?

## Tinjauan Keilmuan

---

- ❑ Teknik tradisional tidak flexible untuk data mentah
  - ❑ Tidak dapat memproses database dalam jumlah yang besar
  - ❑ Memiliki format tipe data yang berbeda
- ❑ Data mining bermanfaat untuk
  - ❑ Klasifikasi atau pengelompokan data
  - ❑ Formasi hipotesis



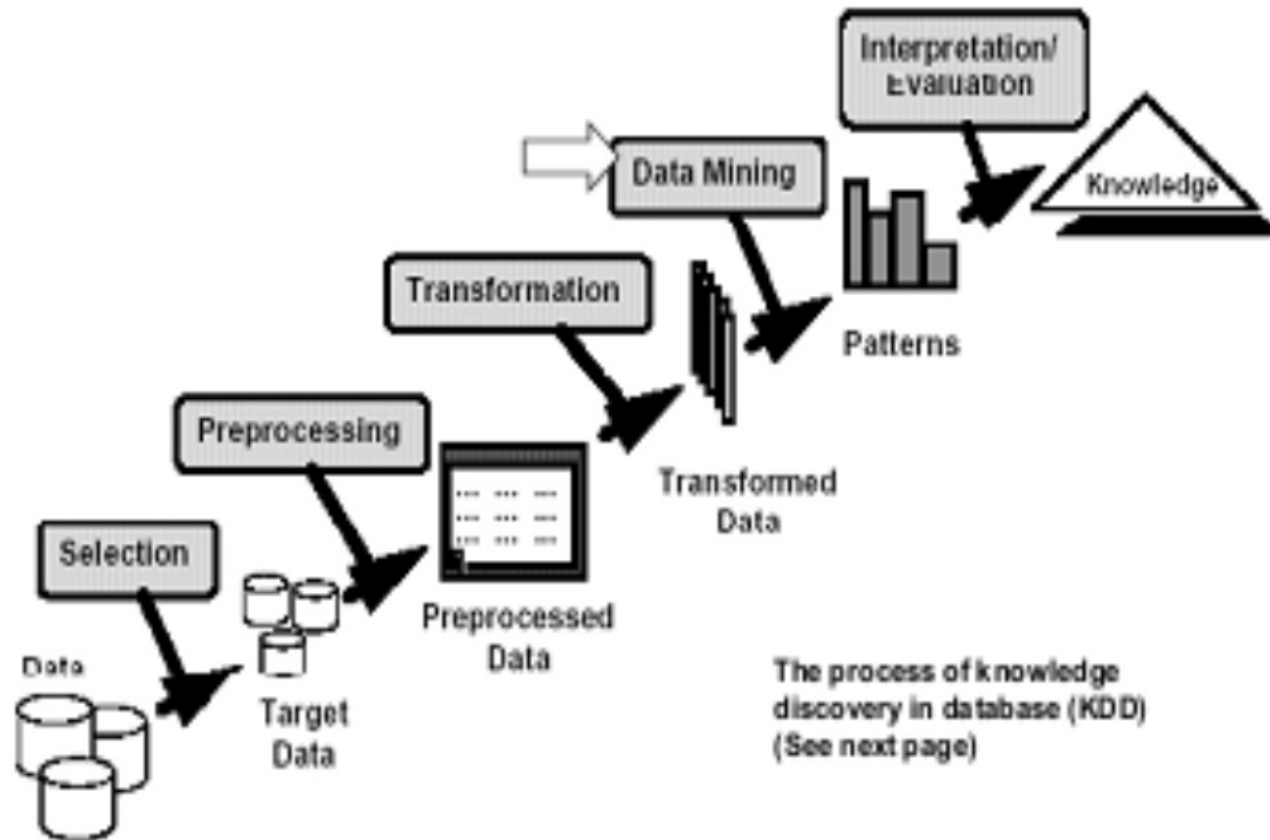
# Apa Itu Data Mining?

---

- ❑ Ekstraksi kompleks yang bersifat implisit terhadap kumpulan data besar yang tidak diketahui sebelumnya namun menyimpan informasi yang potensial dan berharga
- ❑ Ekplorasi dan analisis baik secara otomatis dan semi otomatis dari sejumlah besar data untuk menemukan pola yang berarti

# Proses Penggalian Data

---



# Proses Penggalian Data

---

- ❑ Data mining merupakan bagian yang terintegrasi dari proses penggalian data yang merupakan keseluruhan proses dari proses pengkonversian data dari data mentah ke dalam informasi yang berguna
- ❑ Input data dapat disimpan dalam berbagai file format dan mungkin berada di dalam database pusat

# Proses Penggalian Data

---

## ☐ Preprocessing

- ☐ Menyatukan data dari berbagai sumber
- ☐ Membersihkan data untuk menghilangkan noise dan observasi yang dapat menyebabkan redudansi dan duplikasi hasil
- ☐ pemilihan record data yang diperlukan

## ☐ Post processing dilakukan untuk memperoleh hasil yang diinginkan sesuai dengan DSS (Decision Support System)

# Mengapa Harus Data Mining?

---

- ❑ Informasi berharga sering tersembunyi dalam data yang tidak siap digunakan secara langsung
- ❑ Analisis manual memerlukan waktu lama untuk dapat menemukan informasi berharga didalam data
- ❑ Kebanyakan data bahkan tidak pernah dianalisis sama sekali

# Tantangan Data Mining

---

## ☐ Skalabilitas

- ☐ Memiliki data set dalam jumlah yang sangat banyak
- ☐ Membutuhkan strategi pencarian khusus untuk masing-masing problem

## ☐ Kepemilikan dan distribusi data

- ☐ Secara umum data dimiliki oleh beberapa entitas
- ☐ Untuk mendapatkan hasil maksimal memerlukan informasi dari berbagai sumber

# Tantangan Data Mining

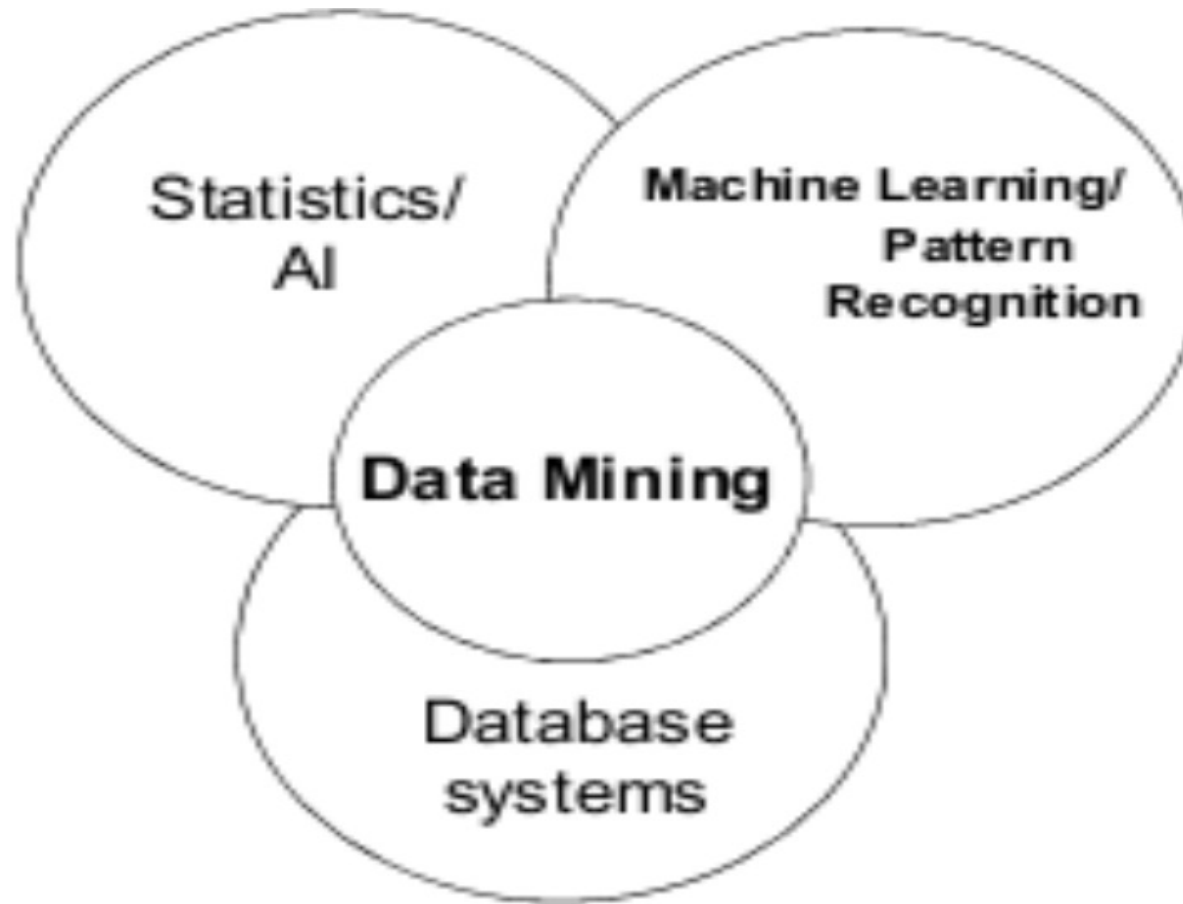
---

- ❑ Heterogenitas dan kompleksitas data
  - ❑ Memerlukan teknik yang dapat handle heterogenitas atribut
  - ❑ Kompleksitas data yang memiliki berbagai macam bentuk



# Origins Data Mining

---



*Terima kasih*

- ARIK SOFAN TOHIR, M.KOM