

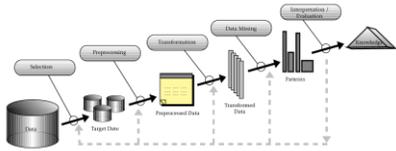
KDD, CRISP-DM, ARSITEKTUR

Hani Irmayanti, M.Kom

Knowledge Discovery in Databases (KDD)

- ▶ Keseluruhan proses untuk mencari dan mengidentifikasi pola (pattern) dalam data, dimana pola yang ditemukan bersifat sah, baru, dapat bermanfaat dan dapat dimengerti.
- ▶ KDD berhubungan dengan teknik integrasi dan penemuan ilmiah, interpretasi dan visualisasi dari pola-pola sejumlah kumpulan data.

Proses KDD



SELECTION

Menciptakan himpunan data target, pemilihan himpunan data, atau memfokuskan pada subset variable atau sampel data, dimana penemuan (discovery) akan dilakukan.

PRE-PROCESSING / CLEANING

- ▶ Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten.
- ▶ Dilakukan proses *enrichment*, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.

TRANSFORMATION

Merupakan proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*.

DATA MINING

- ▶ Pemilihan tugas data mining; pemilihan goal dari proses KDD misalnya klasifikasi, clustering, dll.
- ▶ Pemilihan algoritma data mining untuk pencarian (searching)

Interpretation / Evaluation

Tahap ini merupakan bagian dari proses KDD yang mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

CRISP-DM

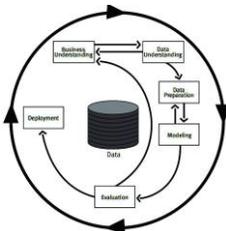
(Cross-Industry Standard Process for Data Mining)

CRISP-DM

(Cross-Industry Standard Process for Data Mining)

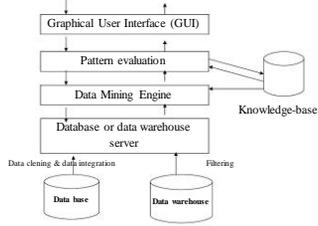
konsorsium perusahaan yang didirikan oleh Komisi Eropa pada tahun 1996 dan telah ditetapkan sebagai proses standar dalam *data mining* yang dapat diaplikasikan di berbagai sektor industri.

Siklus Hidup Pengembangan Data Mining



ARSITEKTUR DATA MINING

Arsitektur : Sistem Data Mining



2

Generik 3-tier arsitektur untuk Data Mining.



SELESAI
TERIMA KASIH