

Konsep Data Mining

Nugroho D.S.

Mengapa Data Mining?

- * Data yang dikumpulkan semakin banyak
- * Teknik penyimpanan data semakin tinggi
- * Proses ekstraksi informasi semakin besar
- * Tekanan kompetisi semakin kuat
- * Analisis data Tradisional tidak dapat digunakan

Apa itu Data Mining?

- * Disiplin ilmu yang mempelajari metode untuk mengekstrak pengetahuan atau menemukan pola dari suatu data
- * Sehingga Data mining sering disebut Knowledge Discovery in Database (KDD)

Konsep Transformasi

DATA

- * Tidak membawa arti, merupakan kumpulan dari fakta-fakta tentang suatu kejadian
- * Suatu catatan terstruktur dari suatu transaksi
- * Merupakan materi penting dalam membentuk informasi

Konsep Transformasi

INFORMASI

- * *Informasi adalah data yang telah diolah menjadi bentuk yang memiliki arti bagi si penerima dan bermanfaat bagi pengambilan keputusan saat ini atau mendatang .
(Raymond Mc.leod)*
- * *mendefinisikan informasi sebagai data yang telah diproses sedemikian rupa sehingga meningkatkan pengetahuan seseorang yang menggunakan data tersebut. (Abdul Kadir)*

Pengetahuan

- * Gabungan dari suatu pengalaman, nilai, informasi kontekstual dan juga pandangan pakar yang memberikan suatu framework untuk mengevaluasi dan menciptakan pengalaman baru dan informasi
(*Thomas H. Davenport, Laurence Prusak*)
- * Bisa berupa solusi pemecahan suatu masalah, petunjuk suatu pekerjaan dan ini bisa ditingkatkan nilainya, dipelajari dan juga bisa diajarkan kepada yang lain

Data - Informasi – Pengetahuan

Data Kehadiran Pegawai

NIP	TGL	DATANG	PULANG
1103	02/12/2004	07:20	15:40
1142	02/12/2004	07:45	15:33
1156	02/12/2004	07:51	16:00
1173	02/12/2004	08:00	15:15
1180	02/12/2004	07:01	16:31
1183	02/12/2004	07:49	17:00

Data - Informasi – Pengetahuan

Informasi Akumulasi Bulanan Kehadiran Pegawai

NIP	Masuk	Alpa	Cuti	Sakit	Telat
1103	22				
1142	18	2		2	
1156	10	1	11		
1173	12	5			5
1180	10			12	

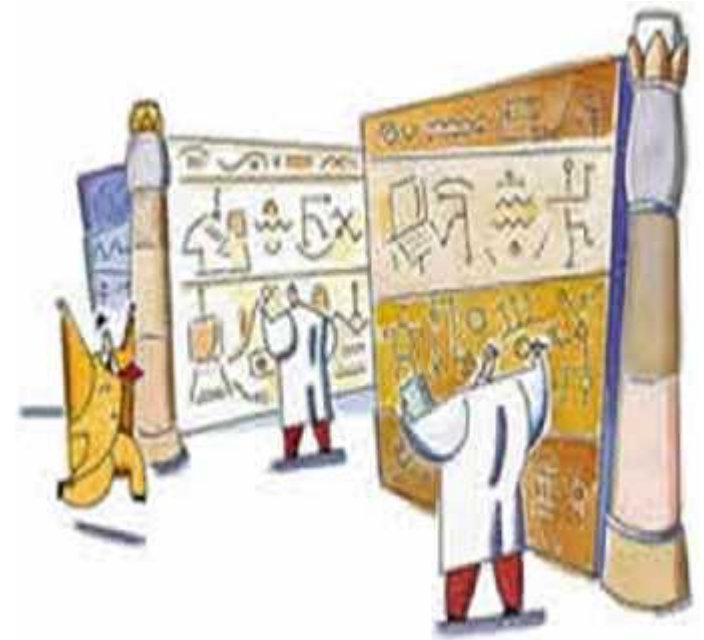
Data - Informasi – Pengetahuan

Informasi Kondisi Kehadiran Mingguan Pegawai

	Senin	Selasa	Rabu	Kamis	Jumat
Terlambat	7	0	1	0	5
Pulang Cepat	0	1	1	1	8
Izin	3	0	0	1	4
Alpa	1	0	2	0	2

Data - Informasi – Pengetahuan

- Pengetahuan tentang kebiasaan pegawai dalam jam datang/pulang kerja
- Pengetahuan tentang bagaimana teknik meningkatkan kehadiran pegawai → kebijakan



Data - Informasi - Pengetahuan - Kebijakan

- * **Kebijakan penataan jam kerja** karyawan khusus untuk hari senin dan jumat
- * Peraturan jam kerja:
 - * Hari Senin dimulai jam 10:00
 - * Hari Jumat diakhiri jam 14:00
 - * Sisa jam kerja dikompensasi ke hari lain:
 1. Senin pulang setelah maghrib (**bayar hutang 2 jam**)
 2. Rabu dan kamis bayar hutang setengah jam di pagi hari dan setengah jam di sore hari (**bayar hutang 2 jam**)

Definisi Data Mining

- * Melakukan **ekstraksi** untuk mendapatkan **informasi penting** yang sifatnya **implisit** dan sebelumnya tidak diketahui, dari suatu data (*Witten et al., 2011*)
- * Kegiatan yang meliputi pengumpulan, pemakaian data historis untuk **menemukan keteraturan, pola dan hubungan** dalam set data berukuran besar (*Santosa, 2007*)

Irisan Bidang Ilmu Data Mining

1. Statistik:

- * Lebih bersifat teori
- * Fokus ke pengujian hipotesis

2. Machine Learning:

- * Lebih bersifat heuristik
- * Fokus pada perbaikan performansi dari suatu teknik learning

3. Data Mining:

- * Gabungan teori dan heuristik
- * Fokus pada seluruh proses penemuan knowledge dan pola
- * Termasuk data clening, learning dan visualisasi hasilnya