

DATA MINING

TUGAS 1



DISUSUN OLEH :

WA ODE ARUM PUTRI WAHIDAH DANI

D121181515

DEPARTEMEN TEKNIK INFORMATIKA

FAKULTAS TEKNIK

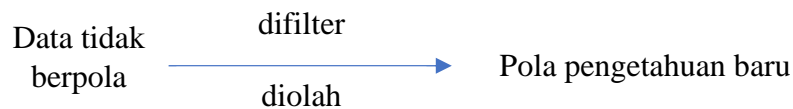
UNIVERSITAS HASANUDDIN

MAKASSAR

2021

Quiz 1

1. Data adalah kumpulan fakta yang terstruktur yang memberikan gambaran yang luas mengenai suatu keadaan. Informasi adalah data atau fakta yang diolah yang memiliki arti. Sedangkan pengetahuan adalah hasil keluaran atau nilai dari informasi.
2. Data mining adalah proses ekstraksi data (sebelumnya tidak diketahui dan dianggap tidak berguna) menjadi suatu pola pengetahuan baru.



3. Peran utama data mining :
 - a. Estimasi
 - b. Prediksi
 - c. Klasifikasi
 - d. Klustering
 - e. Asosiasi
4. Pemanfaatan data mining diberbagai bidang :
 - a. Analisa pasar dan manajemen : Identifikasi kebutuhan customer.
 - b. Analisa Perusahaan dan Manajemen resiko : Persaingan dengan kompetitor sejenis
 - c. Keuangan : Deteksi transaksi-transaksi keuangan yang mencurigakan (seperti money laundry)
5. Dari data diatas didapatkan pengetahuan siswa SMA dengan jenis kelamin laki-laki dapat lulus tepat waktu, nilai data nilai UN dan indeks prestasi tidak begitu memberi pengaruh besar pada tepatnya waktu lulus.

Quiz 2

1. Data mining adalah kemampuan untuk mencari informasi dari data yang dianggap tidak berguna menjadi sebuah informasi yang berharga, dapat dianalogikan dengan proses menambang emas dari tumpukan lumpur.
2. Database yang memodelkan data sebagai fakta, dimensi,dan numerik untuk menganalisis data dalam jumlah besar yang tujuannya untuk mengambil keputusan. Contohnya InterSystems Cathe, ContourCube, dan Cognoa PowerPlay.

Quiz 3

1. 5 Peran utama data mining :
 - a. Estimasi
 - b. Prediksi
 - c. Klasifikasi
 - d. Klustering
2. Estimasi adalah proses untuk menemukan suatu karakteristik penting dari data dalam suatu basis data dengan cara menduga atau menebak dari sebuah nilai yang belum diketahui. Sedangkan prediksi adalah proses untuk menemukan jenis, pola atau nilai dari data dengan memperkirakan hal tersebut yang belum diketahui untuk masa mendatang.
3. Prediksi sama dengan klasifikasi dan estimasi yang membedakannya hanya hasil dalam prediksi yang terjadi dimasa yang akan datang. Sedangkan, klasifikasi adalah proses penemuan model/fungsi yang menjelaskan atau membedakan konsep/kelas data untuk mendeskripsikan data yang penting serta dapat meramalkan kecenderungan data pada masa depan.
4. Klasifikasi adalah yang menggunakan data dengan target/class/label berupa nilai kategorikal (nominal), mode data mining memeriksa set record yang besar, dimana setiap record memiliki informasi variabel target dan set input. Sedangkan klustering adalah pengelompokan data, record, observasi atau kasus ke dalam objek-objek yang mirip dan memiliki perbedaan bila dibandingkan dengan data dari kluster lain. Didalam clustering tidak terdapat variabel/tidak memiliki target/class/label.
5. Klustering adalah pengelompokan data, hasil observasi dan kasus ke dalam class yang mirip antara satu dengan yang lain, dan memiliki perbedaan bila dibandingkan dengan data dari kluster lain. Sedangkan, asosiasi adalah suatu tugas untuk menemukan atribut yang “muncul bersamaan” yang mencoba menemukan aturan untuk menjumlahkan/membandingkan hubungan antara dua atau lebih atribut.
6. Estimasi adalah algoritma yang menggunakan data dengan target/class/label berupa bilangan numerik (kontinyu) dan bukan kategorikal (nominal atau diskrit). Estimasi nilai dari variable target ditentukan berdasarkan nilai dari variabel prediktor (atribut). Sedangkan klasifikasi adalah algoritma yang menggunakan data dengan target/class/label berupa nilai kategorikal (nominal).
7. Estimasi adalah algoritma yang menggunakan data dengan target/class/label berupa bilangan numerik (kontinyu) dan bukan kategorikal (nominal atau diskrit). Estimasi nilai dari variable target ditentukan berdasarkan nilai dari variabel prediktor (atribut). Sedangkan klustering adalah pengelompokkan data, hasil observasi dan kasus ke dalam class yang mirip, klustering tidak memiliki target/class/label.
8. Supervised learning adalah algoritma yang paling sering digunakan dalam dunia data science dibandingkan dengan unsupervised learning. algoritma data mining yang termasuk adalah (estimasi, prediksi, klasifikasi). Perbedaan kedua algoritma tersebut terletak pada

cara belajar untuk membuat suatu prediksi maupun klasifikasi. Dalam supervised learning, algoritma tersebut seolah-olah dilatih terlebih dahulu agar dapat melakukan prediksi maupun klasifikasi. Sedangkan, pada unsupervised learning, untuk melakukan prediksi maupun klasifikasi mereka tidak perlu dilatih terlebih dahulu. Sehingga dapat dikatakan bahwa supervised learning membutuhkan data training agar mampu melakukan prediksi maupun klasifikasi. algoritma data mining yang termasuk adalah klastering.

9. Tahapan utama proses data mining :
 - a. Analisis kebutuhan data : data collecting, data cleaning, data processing, dan data preparation
 - b. Penentuan metode dengan algoritma data mining (proses mining data)
 - c. Dari proses mining data menghasilkan output berupa pola, model, atau knowledge berupa informasi dan pengetahuan
 - d. Evaluasi terkait keakuratan dan nilai error