Sabri Tajuddin

10184043

Si\_RS-6.1

1) Data mining merupakan sebuah proses pengumpulan informasi penting dari sebuah data yang besar

2) Fungsi untuk memahami lebih jauh tentang data yang diamati. Dengan melakukan sebuah proses diharap bisa mengetahui perilaku dari sebuah data

3) Manfaat data mining yang sangat penting dalam pertumbuhan sebuah perusahaan

4) Saat ini data mining tuh sangat diperlukan terutama dalam mengelola data yang sangat besar, tujuannya ialah untuk memudahkan aktifitas recording suatu transaksi dan untuk proses data warehousing agar dapat memberikan informasi yang akurat bagi penggunanya.

5) Berikut 7 proses yang umumnya ada pada Knowledge Discovery in Database (KDD) :

Pembersihan Data (Data Cleaning)

Integrasi Data (Data Integration)

Pemilihan Data (Data Selection)

Transformasi Data (Data Transformation)

Penambangan Data (Data Mining)

Evaluasi Pola (Pattern Evaluation)

6)yakni teknik klasifikasi yang digunakan yaitu Support Vector Machine, AdaBoost, Naïve Bayes, Constant, KNN, Tree, Random Forest, Stochastic Gradient Descent, dan CN2 Rule. Proses klasifikasi diawali dengan preprocessing data untuk melakukan penghilangan missing value dan pemilihan fitur pada dataset. Pada tahap evaluasi digunakan teknik 10 fold cross validation. Setelah dilakukan pengujian, didapatkan bahwa hasil klasifikasi menunjukkan akurasi terbaik diperoleh oleh model Tree, Constant, Naive Bayes, dan Stochastic Gardient Descent

7) yakni menghasilkan sebuah aplikasi dari yang didalamnya terdapat informasi dari hasil analisis teknik data mining

8) Clustering adalah metode penganalisaan data, yang sering dimasukkan sebagai salah satu metode Data Mining, yang tujuannya adalah untuk mengelompokkan data dengan karakteristik yang sama ke suatu 'wilayah' yang sama dan data dengan karakteristik yang berbeda ke 'wilayah' yang lain.