

TUGAS PERTEMUAN 11

Nama : Ahmad Haidar

Npm : 43E57006185006

1.Data mining merupakan proses analisa data untuk menemukan suatu pola dari kumpulan data tersebut. Data mining mampu menganalisa data yang besar menjadi informasi berupa pola yang mempunyai arti bagi pengambil keputusan.Dalam skripsi ini yang akan dibahas adalah mengenai prediksi laju pertumbuhan penduduk pada Badan Pusat Statistik Kabupaten Deli Serdang. Salah satu teknik yang ada pada data mining adalah estimasi. Pada skripsi ini akan dibahas teknik estimasi yang diterapkan untuk menemukan pola yang terjadi pada data penduduk terutama hal yang berkaitan dengan laju pertumbuhan penduduk pada Badan Pusat Statistik Kabupaten Deli Serdang. Teknik estimasi yang akan digunakan adalah Regresi Linear Berganda. Metode ini dipilih karena mampu membuat suatu estimasi/prediksi dengan memanfaatkan data-data lama mengenai laju pertumbuhan penduduk. Sehingga dapat dihasilkan suatu pola hubungan antara atribut-atribut yang mempengaruhi laju pertumbuhan penduduk. Metode Regresi Linear Berganda ini bertujuan untuk membuat persamaan regresi dan prediksi terbaik berdasarkan atribut-atribut yang ada .

2.Badan Pusat Statistik (BPS) adalah lembaga pemerintah nonkementrian yang bertanggung jawab langsung kepada presiden. Salah satu lembaga BPS yang ada di Sumatera Utara adalah BPS Deli Serdang. Termasuk salah satu tugas dan fungsi BPS adalah melakukan pendataan statistik terhadap jumlah penduduk. Adapun pendataan penduduk rutin oleh dilakukan BPS Deli Serdang setiap beberapa tahun sekali. Namun BPS Deli Serdang mengalami kesulitan dalam mengestimasi laju pertumbuhan penduduk setiap tahunnya dikarenakan beberapa kendala. Dari beberapa karya ilmiah ternyata ada suatu cara untuk dapat mengestimasi laju pertumbuhan penduduk yaitu dengan menggunakan teknik Data Mining. Data Mining merupakan penambangan atau penemuan informasi baru dengan mencari pola atau aturan tertentu dari sejumlah data dalam jumlah besar yang diharapkan dapat mengatasi kondisi tersebut. Data Mining sendiri memiliki beberapa metode diantaranya adalah metode regresi linear berganda. Kemudian metode regresi linear berganda akan digunakan untuk membangun suatu aplikasi perangkat lunak yang berbasis desktop programming. Yang diharapkan dapat membantu memecahkan masalah dalam mengestimasi laju pertumbuhan penduduk. Kehadiran data mining dilatarbelakangi dengan masalah data explosion yang dialami akhir-akhir ini dimana banyak perusahaan/bank/organisasi telah mengumpulkan data sekian tahun lamanya (data pembelian, data penjualan, data nasabah, data transaksi, dan lain-lain). Seperti yang terjadi pada sebuah perusahaan asuransi, data yang tersimpan selama ini hanya sebagai dokumentasi dan hanya dipakai untuk kebutuhan transaksi saja. Pertanyaannya sekarang, apakah data tersebut akan dibiarkan menggunung, tidak berguna lalu dibuang, Metode Regresi Linear

Regresi linear merupakan satu cara prediksi yang menggunakan garis lurus untuk menggambarkan hubungan

diantara dua variabel atau lebih. Variabel adalah besaran yang berubah-ubah nilainya. Selanjutnya variabel

tersebut terbagi atas dua jenis yaitu variabel pemberi pengaruh dan variabel terpengaruh. Variabel pemberi

pengaruh dapat dianalogikan sebab, sementara variabel terpengaruh merupakan akibat (Sani Susanto, 2010 :

33).

Informasi yang dihasilkan dari data mining dengan metode Regresi Linear ini sendiri juga bisa dijadikan

sebagai dasar untuk pengambilan keputusan. Contoh kasus metode data mining ini biasanya dipakai dalam

data perhitungan sebuah waktu yang dibutuhkan seseorang untuk mengantarkan pizza kepada konsumennya,

yang nantinya akan menunjukkan jarak yang akan ditempuh oleh pengantar pizza itu, jumlah pesanan yang

dipesan oleh konsumen dan juga banyaknya lampu lalu lintas yang akan dilewatinya.

Algoritma Regresi Linear termasuk jenis aturan classification and regression pada data mining, selain Regresi

Linear yang termasuk pada golongan ini adalah Support Vector Machine, Logistic Regression dan lain-lain.

Analisis Regresi Linear adalah teknik data mining untuk menentukan bahwa terdapat hubungan antara

variabel yang ingin diramalkan (variabel tak bebas) dengan variabel lain (variabel bebas).

Selanjutnya

peramalan ini didasarkan pada asumsi bahwa pola pertumbuhan data historis yang bersifat linier, walaupun

sebenarnya tidak 100% linier. Pola pertumbuhan ini didekati dengan suatu model yang menggambarkan

hubungan-hubungan yang terkait dalam suatu keadaan.

Algoritma Regresi Linear Berganda

Uji coba dan analisa adalah untuk melihat bagaimana suatu proses pada software data mining, tujuan dari

proses dan output yang diharapkan. Ujicoba ini bertujuan untuk melakukan perhitungan algoritma untuk

mencari nilai prediksi terbaik yang akan menjadi pembagi terhadap atribut yang lainnya, dengan menggunakan perhitungan (rumus) regresi linear berganda. Regresi linear berganda akan dilatihkan terhadap

pola data input dan target sebenarnya.

Regresi linier merupakan bentuk hubungan di mana variabel bebas X maupun variabel tergantung Y sebagai faktor yang berpangkat dua. Persamaan umum algoritma regresi linier berganda, adalah :

$$Y = a_0 + a_1.X_1 + a_2.X_2 + \dots + a_n.X_n$$

Keterangan :

Y = variabel terikat / variabel dependent (nilai yang diprediksi)

a_0, a_1, a_2, a_n = koefisien regresi

X_1, X_2, \dots, X_n = variabel bebas / variabel independent.

Analisis regresi linier berganda digunakan untuk mengestimasi pertumbuhan penduduk sebagai (Y) di BPS Kabupaten Deli Serdang ditinjau dari 2 variabel yaitu Jumlah Laki-Laki (X_1) dan Jumlah Perempuan

(sebagai X_2) yang akan dilakukan prediksi menggunakan analisis regresi linier berganda.