

Dataset

Data adalah fakta dan angka (dapat juga disebut sebagai data mentah) yang berhubungan dengan konteks suatu permasalahan, Data terdiri dari dua aspek yaitu Object dan atribut, contoh object manusia, pohon, binatang, contoh atribut misalnya umur, tinggibadan, Beratbadan.

Dataset merupakan kumpulan objek data. Dataset memiliki nama lain record,point,vector,pattern,event,observasi,case atau data. Object data digambarkan dengan menggunakan sejumlah atribut yang menangkap karakteristik dari object data tersebut. Atribut disebut juga sebagai karakteristik,varibel,field,fitur,atau dimensi.

Atribut

Atribut dataset (atau biasanya disebut "fitur" dalam konteks machine learning) adalah variabel yang digunakan untuk menggambarkan atau mengukur sebuah objek atau entitas dalam dataset. Atribut dataset dapat berupa angka, teks, atau nilai diskrit lainnya, dan setiap atribut memiliki satu atau lebih nilai untuk setiap entitas dalam dataset.

Ini adalah contoh dataset dari Daerah Jawa Timur

Anggota perpustakaan di provinsi jawa timur adalah masyarakat Jawa Timur dan non-Jawa Timur yang memiliki MoU dengan Dinas Perpustakaan dan Kearsipan Provinsi Jawa Timur

periode_update	anggota_laki_laki	anggota_perempuan
Januari 2022	234,600	351,900
Februari 2022	250,090	375,135
Maret 2022	263,288	394,933
April 2022	282,184	423,276
Mei 2022	289,800	434,700
Juni 2022	281,027	421,541
Juli 2022	223,550	335,325
Agustus 2022	272,139	408,208
September 2022	279,380	419,070
Oktober 2022	308,223	462,334
November 2022	299,481	449,222
Desember 2022	311,941	467,912

Ini adalah grafik dari Dataset di atas

