

Penjelasan

Dataset yang diberikan merupakan record pasien yang akan melakukan konsultasi pada suatu klinik X. Sering ada kejadian, calon pasien yang telah mendaftarkan diri untuk konsultasi atau berobat tidak hadir pada hari-H yang telah ditentukan.

Silahkan buat model untuk menentukan variabel Label akan bernilai 0 atau 1 pada data train. Dari model yang didapat, terapkan pada data test untuk memprediksi nilai Label dari calon pasien.

Beri nama file hasil prediksi dengan format "DataMining_Prediksi_NamaTim.csv"
Beri nama file paper dengan format "DataMining_Paper_NamaTim.pdf"

Upload hasil prediksi dan paper sesuai format pada [web C-Compiler 2017](http://web.c-compiler.com) sebelum 28 April 2017 pukul 23.59.

Deskripsi File

- train.csv : Data Train berupa 299701 data dengan 27 variabel
- test.csv : Data Test berupa 300 data dengan 26 variabel
- sampleSubmission.csv : contoh file hasil prediksi
- Format_Paper.docx : contoh format isi paper

Data Fields

- Usia : Usia calon pasien
- Jenis_Kelamin : Jenis kelamin calon pasien
- Waktu_Pendaftaran : Tanggal calon pasien melakukan registrasi
- Waktu_Konsultasi : Tanggal pasien akan melakukan konsultasi
- Hari : Nama hari dari Waktu_Konsultasi
- Dis_1-Dis_17 : Suatu penyakit yang akan dikonsultasikan
- Subsidi : Label apakah calon pasien mendapat subsidi(1) atau tidak(0)
- Sms_Reminder : Reminder SMS ke calon pasien untuk konsultasi
- Salutation : Salam untuk calon pasien
- Label : Label apakah calon pasien datang(1) atau tidak(0)

Penilaian

Penilaian penyisihan dilihat dari akurasi prediksi dan konten paper. Penilaian akurasi digunakan,
$$\text{Akurasi} = \frac{tp+tn}{tp+tn+fp+fn}$$

Apabila masih ada yang kurang jelas bisa ditanyakan ke KEN (sms 085740106193/ id line kurniawanen)