

Prediksi Hasil Income Menggunakan Metode Naive Bayes

Bintang Peryoga

IF-40-04

Fakultas Informatika, Universitas Telkom

Jl. Telekomunikasi no.1 Terusan Buah Batu, Kab. Bandung, Jawa Barat

bintangperyoga.bp@gmail.com

I. PENDAHULUAN

Laporan ini berisi cara penentuan hasil *income* menggunakan metode Naïve Bayes. Rumus Naïve Bayes:

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)}$$

Dengan menggunakan metode ini, diharapkan dapat membantu menentukan hasil *income* dan mengklasifikasi mana saja yang termasuk golongan penghasilannya '>50K' atau '<=50K'.

II. PEMBAHASAN

A. Analisis Masalah

Terdapat 160 *data train* yang memiliki 8 atribut (Age, workclass, education, marital-status, occupation, relationship, hours-per-week, dan income) dan setiap atribut memiliki tipenya masing-masing. Programmer diminta untuk menentukan nilai atribut *income* dari *data test* sebanyak 40 data dengan menggunakan metode Naïve Bayes.

B. Strategi Penyelesaian Masalah

i. Menentukan *library*

Saya hanya menggunakan *pandas*

ii. Membuat *function*

Saya membuat 2 fungsi(count dan naïve bayes) dimana fungsi count untuk menentukan berapa banyak data tersebut muncul dan fungsi naïvebayes untuk menghitung peluang

iii. Mengimpor data excel

Saya mengimpor dengan cara `pd.read_excel('nama file')`

iv. Ada 4 langkah pengerjaan, yaitu:

- Menghitung peluang income
- Menghitung peluang dari masing-masing atribut
- Menentukan hasil prediksi, setelah itu dimasukkan ke array `datatest_income`
- Masukkan isi dari array tersebut ke file `datatest` lalu export ke file `TebakanTugas1.xlsx`

Ket: codingan terlampir

III. KESIMPULAN

Kesimpulan dari laporan ini, 40 data telah berhasil diklasifikasikan berdasarkan metode naïve bayes apakah data tersebut termasuk ke kelas income '>50K' atau kelas income '<=50K'. Dengan menggunakan metode naïve bayes ini programmer dapat dengan mudah menentukan prediksi hasil yang dicari.

