

Nama : Gigas Taufan Arvyanto

NIM : 1301164211

Kelas : IF-40-12

TUGAS 4 Machine Learning

1. Identifikasi Masalah

Diberikan dua buah file di mana masing-masing digunakan sebagai data train dan data test. Data train terdiri dari 298 data yang setiap barisnya terdiri dari 3 atribut yaitu, x1, x2, dan class. Class merupakan nilai kelas atau golongan (1 atau 2). Untuk data test terdiri dari 75 data. Pada data test sudah terdapat atribut x1 dan x2. Dengan menggunakan Ensemble Learning dengan metode Bagging dilakukan penentuan kelas untuk setiap data pada data test tersebut.

2. Desain Program

Untuk menyelesaikan permasalahan ini digunakan metode Bootstrap Aggregation (Bagging). Metode ini menggunakan Naïve Bayes untuk melakukan klasifikasi. Berikut merupakan langkah-langkah yang dilakukan penulis untuk menyelesaikan permasalahan yang diberikan:

a. Membuka File

Membuka file “TrainsetTugas4ML.csv” sebagai datatrain, dan “TestsetTugas4ML.csv” sebagai datatest. Datatrain telah memiliki nilai Class untuk datanya, sedangkan datatest belum memiliki nilai Class pada datanya.

b. Membuat Bootstrap

Membuat himpunan yang disebut dengan bootstrap. Himpunan bootstrap diisi sebanyak 63% dari seluruh datatrain yang dipilih secara random. Kemudian mencari nilai Standar Deviasi dari data tersebut.

c. Menghitung Naïve Bayes

Menggunakan rumus Naïve Bayes untuk menemukan kelas dari dataset.

$$P(x_k|C_i) = \frac{1}{\sigma_{ik} \sqrt{2\pi}} e^{-\frac{(x_k - \mu_{ik})^2}{2\sigma_{ik}^2}}$$

$$\pi = 3,1416 \quad e = 2,7183$$

d. Membuat model

Ulangi langkah b-d sebanyak beberapa kali, lakukan perulangan sebanyak angka ganjil.

e. Hitung Jumlah

Setelah terdapat data hasil sebanyak perulangan tadi lalu bandingkan. Setiap data di dalam model telah memiliki kelas, kemudian untuk setiap data hitung jumlah masing-masing kelas. Kelas dengan jumlah terbanyak ditunjuk sebagai kelas akhir dari data tersebut.

3. Hasil Program

Berikut merupakan hasil dari program yang dibangun:

1	2	24	1	47	1		
2	2	25	1	48	1		
3	2	26	1	49	1		
4	2	27	1	50	1		
5	2	28	2	51	1		
6	2	29	1	52	1		
7	2	30	2	53	1		
8	2	31	2	54	1		
9	2	32	1	55	1		
10	2	33	2	56	1		
11	2	34	1	57	1		
12	2	35	1	58	1		
13	2	36	2	59	1		
14	2	37	1	60	1		
15	2	38	1	61	1		
16	1	39	1	62	1		
17	2	40	1	63	1		
18	2	41	1	64	1	70	1
19	2	42	2	65	1	71	1
20	1	43	1	66	1	72	1
21	2	44	1	67	1	73	1
22	1	45	1	68	1	74	1
23	2	46	1	69	1	75	1