

1. Deskripsi Masalah

Ensamble Learning adalah salah satu teknik learning yang memiliki beberapa model yang digunakan, yaitu *bagging*, *boosting* dan *random forest*. Bagging (Bootstrap aggregating) ditemukan oleh Leo Breiman pada 1994. Merupakan metode ensemble learning efektif yang pertama kali diusulkan. Bagging dikenal sebagai salah satu metode yang paling simpel dalam arching (adaptive reweighting and combining), sebuah terminologi umum yang mengacu pada penggunaan kembali atau pemilihan data untuk meningkatkan akurasi klasifikasi (Sewell 2018). Berikut adalah algoritma dari metode *Bagging*:

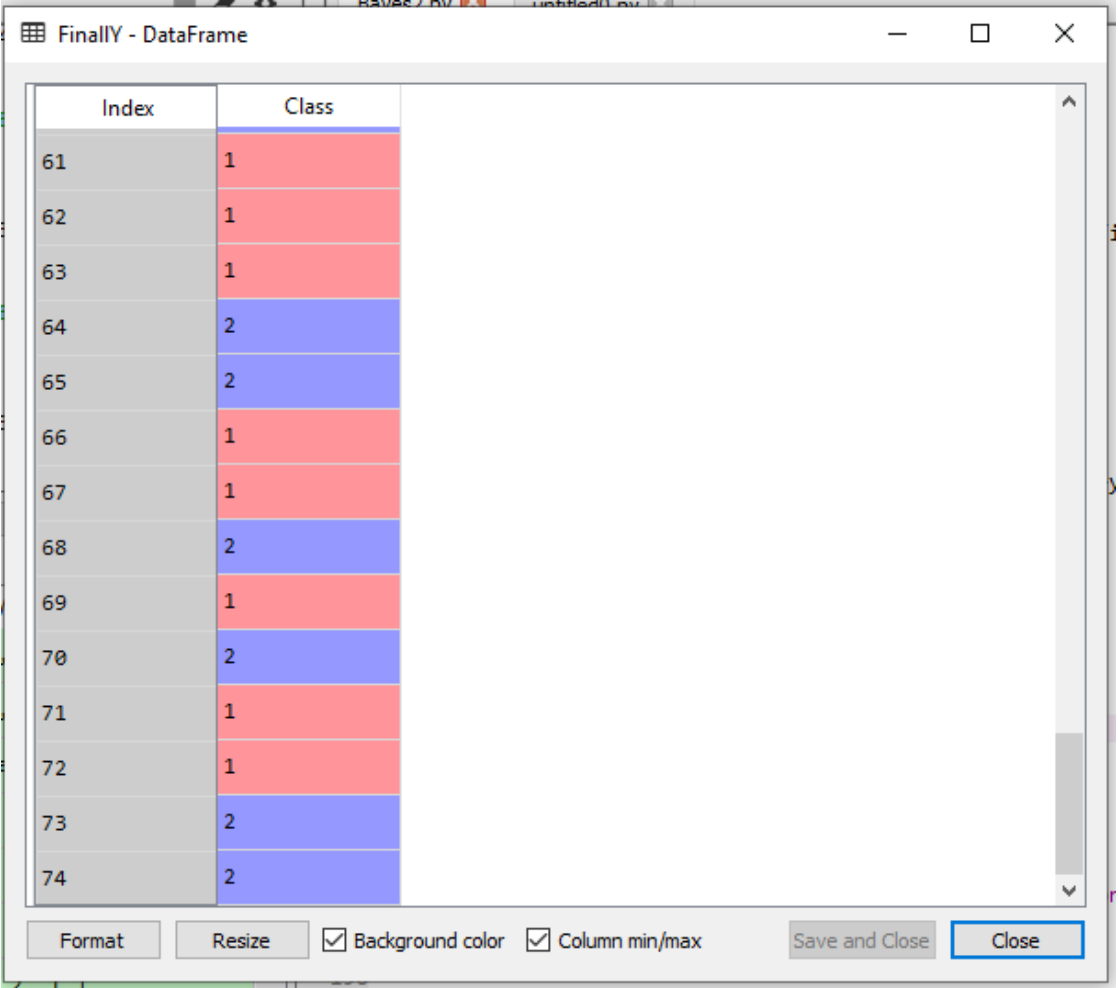
1. Membagi Data Train menjadi beberapa sampel data
 - Misal : membuat 3 sampel data dimana satu sampel data berisi 100 data
 - Cara membagi isi dari sampel data tersebut adalah dengan merandom nilai
 - 1 sampel data bisa jadi memiliki anggota yang duplikat
 - 1 sampel data bisa jadi sama dengan sampel yang lain
2. Membuat model dari sampel data yang telah dibuat
3. Melakukan training model terhadap Data Train
 - Melakukan prediksi nilai kelas dari data train ke masing masing model yang dibuat dari data sampel
4. Membuat list berisi sum(jumlah) dari nilai kelas data train masing masing model
5. Melakukan penandaan dari nilai sum yang didapatkan
 - Misal : if $\text{sum} < 2$ then $\text{sign} = 1$, else : $\text{sign} = 2$
6. Membuat list berisi sign yang telah ditentukan
7. Menghitung akurasi dari nilai kelas Data Train dengan list Sign yang telah dibuat
8. Melakukan step ke 3 sampai ke 5 terhadap DataTest

2. Analisis

Diberikan dua buah data csv yaitu TrainsetTugas4ML.csv yang berisi 297 data dengan 2 fitur (X1 dan X2) serta 1 kolom kelas dan TestsetTugas4ML.csv yang berisi 2fitur(X1 dan X2) tanpa kolom kelas . Dari dua buah data tersebut kita diminta untuk melakukan klasifikasi terhadap TestsetTugas4ML.csv untuk menemukan kelasnya dengan metode *Bagging*, yaitu salah satu teknik dari *Ensamble Learning*. Setelah menemukan kelas dari 75 data TestsetTugas4ML.csv. Kelas yang bisa diberikan pada DataTest adalah 1 atau 2.

2.1 Algoritma

Untuk menjawab persoalan kali ini, saya menggunakan Bahasa python untuk menyelesaikannya, saya juga menggunakan beberapa library untuk memudahkan pencarian model. Saya menggunakan library pandas, numpy serta scikitlearn untuk membantu saya dalam menyelesaikan permasalahan. Berikut adalah hasil dari pemberian kelas dari data test



Index	Class
61	1
62	1
63	1
64	2
65	2
66	1
67	1
68	2
69	1
70	2
71	1
72	1
73	2
74	2