

Tugas 3: Ensemble Learning (Bagging & Random Forest)

MK Pembelajaran Mesin

Semester Ganjil 2022/2023

Ketentuan Umum:

1. Tugas dilakukan secara **berkelompok**
 - a. 3-4 orang untuk kelas regular/INT
 - b. 3 orang untuk kelas PJJ
2. Waktu pengerjaan selama jam perkuliahan (3 jam).
3. Tidak diperbolehkan menggunakan library untuk bagging atau Random Forest.
4. Library untuk *Decision Tree* classifier diperbolehkan.

Deskripsi Tugas:

Pada tugas ini, anda diminta untuk membangun *Random Forest* classifier terhadap dataset biji-bijian (*seeds_dataset*) dengan menggunakan metode *Bagging*, yang bertujuan untuk mengukur akurasi klasifikasi pada *seeds* dataset. Dataset ini mengklasifikasikan biji menjadi 3 kategori (tipe 1, 2, dan 3) tergantung dari berbagai macam atribut.

File-file tugas dapat diunduh dari: <https://rb.gy/gkzhb3>

Gunakan akun SSO untuk download file

Langkah Pengerjaan Tugas:

1. Gunakan data yang telah disediakan **seeds_dataset.csv**
2. Gunakan template code yang tersedia (**tugas3_seed_template.ipynb**).
3. Pastikan anda menggunakan template ini karena beberapa langkah pembangunan Random Forest yang telah disederhanakan.
4. Rename **tugas3_seed_template.ipynb** menjadi **tugas3_seed_<NIM>.ipynb**, gunakan NIM salah satu anggota kelompok
5. Ikuti langkah-langkah pengerjaan pada template, anda diperbolehkan untuk membuat fungsi sendiri di luar kerangka yang disediakan.
6. Isi excel file **output_<NIM>_random_forest.xlsx** menggunakan angka yang diperoleh dari kode.

File yang dikumpulkan:

1. tugas3_seed_[NIM].ipynb
2. output_[NIM]_random_Forest.xlsx
3. Zip kedua file tersebut. Pengumpulan ke LMS dengan penamaan file **Tugas-3_<NIM>_<Nama>.zip**

** Cukup mengumpulkan 1 saja per kelompok. Gunakan NIM salah satu anggota untuk penamaan file. Isi Nama dan NIM seluruh anggota pada code dan excel file.

Tugas 3								
Tree No.	Selected attributes	Accuracy (on test data)	nr of OOB items	OOB score	No	NIM	Nama	
1					1			
2					2			
3					3			
4					4			
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
		Test accuracy (using Voting)						

Kriteria penilaian:

1. Kelengkapan kode: 25
2. Bootstrapping, seleksi atribut dan baris data (*bagging*): 20
3. Perhitungan OOB dan akurasi: 20
4. Kode voting: 20
5. Jawab pertanyaan di excel file: 15

Tambahan

Contoh program untuk membangun *decision tree* sederhana dan visualisasi sudah disediakan. Bisa dilihat pada file **tugas3_seed_example.ipynb**.