

Assessment CLO-3 Part-1 (atau disingkat dengan Tugas3)  
Ensemble Learning Bagging & Random Forest

Capaian Pembelajaran Mata Kuliah (*Course Learning Outcome*) yang akan dicapai

PLO 5	Mampu menerapkan metode computing dalam pengembangan sistem dan mesin berintelegensia.
CLO 3	Mahasiswa mampu menjelaskan, mengimplementasikan, menganalisis, dan mendesain teknik Ensemble Learning

Pada assessment ini, anda diminta untuk membangun Random Forest classifier terhadap suatu data set dengan menggunakan Bagging. Assesment ini dilakukan secara berkelompok (2-3 orang) dengan pemilihan kelompok disesuaikan dengan pilihan mahasiswa masing-masing.

Petunjuk umum:

- Gunakan data yang telah disediakan **seeds\_dataset.txt**
- Di dalam python code (**tugas3\_seed\_template.ipynb**), **gunakan NIM** salah satu anggota kelompok.
- Tidak diperbolehkan menggunakan library untuk bagging atau Random Forest.
- Library untuk Decision Tree classifier diperbolehkan.
- Contoh program untuk membangun decision tree sederhana dan visualisasi sudah disediakan (**tugas3\_seed\_example.ipynb**).
- Template struktur pengerjaan telah disediakan, tidak harus dipakai, boleh menggunakan kode sendiri.
- Pengumpulan berupa kode (.py atau .ipynb) dan hasil di excel file (**Result\_NIM\_random\_forest.xlsx**)

Tree No.	Columns	Test accuracy	OOB error	No	NIM	Nama
1				1		
2				2		
3				3		
4				4		
5				5		
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
		Test accuracy (using Voting)				

Secara garis besar, assessment ini bertujuan untuk mengukur akurasi Random Forest classifier pada seeds dataset.

1. Dataset tersedia pada file seeds\_dataset.txt
2. Bangun 20 Decision Tree Classifier (boleh menggunakan library)
3. Data yang digunakan untuk membangun decision tree, dipersiapkan menggunakan Bagging + random feature selection
4. Untuk mempermudah, random feature selection hanya dilakukan di awal pembuatan decision tree, tidak pada tiap pemilihan node
5. Out of bag error hanya diukur pada decision tree yang dibangun dengan bootstrap data tersebut, tidak diukur secara keseluruhan Random Forest.