

## TUGAS PENGGANTI ABSEN

Catatan :

Tugas pengganti absen ini di gunakan jika mahasiswa memberikan alasan jelas, misalnya :

1. Ikut turnamen
2. Sakit
3. Acara keluarga : pernikahan, meninggal dsb.
4. Dsb

Tidak berlaku jika alasan :

1. Bangun kesiangian
2. Tidak izin
3. Acara yang tidak jelas / logis.

**A. Tidak hadir 1 – 2x** : meringkas materi pertemuan 1 – 5 baik dari modul maupun dari PPT yang sudah di *upload* ke *Google Classroom*. Di kerjakan tulis tangan dan hasil pekerjaan di scan dan dikumpulkan dalam bentuk PDF.

**B. Lebih dari 2x** : Kerjakan 2 tugas dibawah ini

### DATASET

1. Cari dataset di *UCI Machine Learning Repository*. Bebas. Perhatikan default tasknya apakah *classification* atau *clustering*. (Cara melihatnya sudah saya lampirkan caranya dibawah). Data minimal berjumlah 500 instances/data.
2. Lakukan *pre-processing* data seperti yang telah dilakukan di praktikum menggunakan Notepad dan Excel CSV. Tutorial sudah dicantumkan pada modul praktikum pertemuan 7 dan 8.
3. Simpan data Anda di tempat yang mudah di ingat.

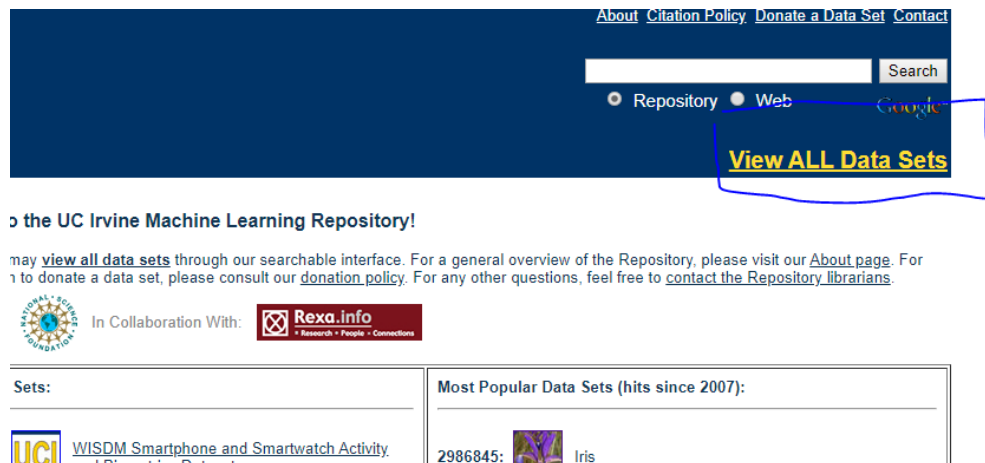
### DATA MINING USING WEKA

1. Gunakan data yang sudah kalian cari di pada tugas dataset diatas.
2. Lakukan *processing data* dengan WEKA. Jika :
  - a. Tipe data *default task* nya adalah *classification* maka lakukan proses klasifikasi dengan metode :
    - 1) Tree
      - RandomForest, atau
      - J48
    - 2) Bayes
      - NaiveBayes, atau
      - BayesNET
  - b. Tipe data *default task* nya adalah *clustering* maka lakukan proses klastering dengan metode :
    - 1) EM
    - 2) SimpleKMeans

3. *Screenshot* tiap prosesnya, dan
4. Jika sudah maka berikan kesimpulan dari model yang telah didapat.
5. Pekerjaan dikumpulkan dalam format PDF. Format file : nama\_npm\_defaulttask  
Contoh : **gustian\_231\_classification**
6. Sertakan juga data CSV yang kalian gunakan.

Bagaimana mengetahui tipe *default task* data kita ?

1. Kunjungi UCI Machine Learning Repository <https://archive.ics.uci.edu/ml/index.php>
2. Lalu klik **View All Dataset**



3. Maka tampilan akan seperti dibawah ini. Tinggal lihat default tasknya apakah klasifikasi atau klustering.

Noted : Data dengan *default task classification* dapat dilakukan proses *classification* dan *clustering*. Data dengan *default task clustering* hanya bisa dilakukan proses *clustering* saja.

UCI Machine Learning Repository		Browse Through: 488 Data Sets					
Machine Learning Repository		Center for Machine Learning and Intelligent Systems					
UCI		View ALL Data Sets					
UCI		Table View List View					
Default Task		Name					
Attribute Type		Data Types					
Data Type		Default Task					
Area		Attribute Types					
		# Instances					
		# Attributes					
		Year					
Classification (360)		Abalone					
Regression (107)		Adult					
Clustering (90)		Annealing					
Other (55)		Anonymous Microsoft Web Data					
Categorical (38)		Arrhythmia					
Numerical (325)		Artificial Characters					
Mixed (55)							
Multivariate (374)							
Univariate (24)							
Sequential (51)							
Time-Series (99)							
Text (55)							
Domain-Theory (23)							
Other (21)							
Life Sciences (110)							
Physical Sciences (52)							
CS / Engineering (178)							
Social Sciences (26)							
Business (32)							
Game (10)							

