

Deskripsi Tugas Besar MK Pembelajaran Mesin CSH3L3 Pembelajaran Mesin Genap 2019/2020

1. Tugas dikerjakan secara individu
2. Harus mematuhi etika plagiasi baik ide, kode, maupun laporan. Segala bentuk pelanggaran akan diberikan sanksi berat, minimal E untuk mata kuliah ini
3. Deadline pengumpulan tugas adalah 13 April 2020, pukul 20.00.
4. Presentasi individu ke dosen dilakukan pada minggu 2 minggu terakhir perkuliahan.
5. Tugas dikumpulkan ke **classroom**.

Deskripsi Tugas:

1. Terdapat 3 buah dataset, setiap mahasiswa **diberikan 1 buah** dataset sesuai table di bagian akhir deskripsi ini. Mahasiswa dibebaskan untuk memilih feature dan label yang akan digunakan.
2. Setiap mahasiswa **harus** mengerjakan **2 buah task** terhadap dataset tersebut, yaitu **Clustering** dan **Classification**.
3. Untuk setiap task, mahasiswa diharuskan setidaknya melakukan hal-hal berikut:
 - a. Formulasi Masalah: mengacu pada 4 tahap From Problem to ML Solution
 - b. Eksplorasi dan Persiapan Data (termasuk split data): Setidaknya menggunakan 2 buah teknik eksplorasi dan/atau penyiapan data. Jelaskan kenapa menggunakan teknik tersebut.
 - c. Pemodelan: Pilih satu algoritma yang akan digunakan. Jelaskan alasan pemilihan tersebut. Buat model menggunakan data hasil proses 3.b. Kerapian dan kejelasan kode sangat penting. Jelaskan dengan detil setiap langkahnya.
 - d. Eksperimen: Wajib melakukan setidaknya 2 eksperimen (menghasilkan 2 model berbeda), misal dengan membedakan fitur yang digunakan, Teknik penyiapan data yang berbeda, algoritma berbeda, dan sebagainya.
 - e. Evaluasi: Pilih metode evaluasi yang ingin digunakan serta jelaskan alasannya. Lakukan evaluasi terhadap model yang digunakan. Berikan analisis terhadap hasil evaluasi.
 - f. Kesimpulan: Sampaikan kesimpulan dari semua proses yang dijalankan, apa saja lesson learned, dan apa saran untuk improvement ke depan.

4. Untuk Task Clustering, mahasiswa **tidak boleh menggunakan library** untuk proses pemodelan (harus coding from scratch).
5. Bahasa pemrograman yang digunakan adalah **Python**.
6. Delivery tugas berupa:
 - a. Laporan diketik dalam A4 menggunakan times new roman size 12, margin 3cm, space 1.15 (.pdf)
 - b. Kode-kode yang siap dieksekusi (.py/.ipynb)
 - c. Dataset hasil eksplorasi
7. Pengumpulan dalam bentuk **.zip/.rar** dengan format filename adalah NIM
 - a. Contoh: 1301160001.zip atau 1301160001.rar
 - b. Pengumpulan selain itu tidak diterima

Penilaian:

1. Laporan (30%):
 - a. Sesuai Format (5%)
 - b. Penggunaan Bahasa yang baik dan benar (5%)
 - c. Kelengkapan (5%)
 - d. Kesesuaian laporan terhadap yang dikerjakan/kode (15%)
2. Kode (40%):
 - a. Kelengkapan semua tahapan (10%)
 - b. Eksekusi kode (20%)
 - c. Kerapian kode, penggunaan fungsi secara pas, pemberian komentar (10%)
3. Presentasi (30%):
 - a. Organisasi konten presentasi (5%)
 - b. Kemampuan menyampaikan presentasi (5%)
 - c. Pemahaman tentang project (20%)
4. Bonus untuk extraordinary results (20%):
 - a. Melakukan lebih banyak eksperimen yang menarik. Pembuatan kode dengan dokumentasi yang sangat lengkap. Penguasaan Teknik-teknik pemrograman advanced.
 - b. Penyajian laporan yang menarik, penyertaan ilustrasi yang membantu pemahaman pembaca, analisis yang tajam, usulan yang bermanfaat.
 - c. Presentasi yang menarik dan orisinal.

Dataset:

Dataset dapat diunduh menggunakan akun **Telkom University** di: <http://bit.ly/project-ml-dataset>

Bandung, 16 Maret 2020

(Tim Dosen Pembelajaran Mesin)