

Deskripsi Tugas Besar Matakuliah Pembelajaran Mesin Semester Ganjil 2021-2022

Ketentuan Umum:

1. Tugas besar dikerjakan dalam dua tahap.
 - a. Tahap pertama dilakukan secara **individual**,
 - b. Sedangkan tahap kedua dilaksanakan secara **berkelompok maksimal dua orang**.
2. Harus mematuhi etika plagiasi, baik ide, kode, maupun laporan. Segala bentuk pelanggaran akan diberikan sanksi berat, minimal E untuk mata kuliah pembelajaran mesin
3. Batas akhir (*deadline*) pengumpulan tugas tahap pertama adalah 13 November 2021 pukul 20.00 WIB sedangkan tahap kedua adalah 04 Desember 2021 pukul 20.00 WIB.
4. Untuk tahap pertama, laporan tugas dan presentasi dilakukan secara individu melalui video presentasi. Sementara itu, untuk tahap kedua, laporan tugas dan presentasi dilakukan pada minggu ke-12, 13, dan 14 di depan dosen pengampu.
5. Tugas dikumpulkan melalui LMS pada masing-masing kelas.

Deskripsi Tugas:

1. Setiap mahasiswa harus mengerjakan dua *task* (**clustering** dan **classification**) terhadap dataset pada link gdrive: <https://drive.google.com/drive/folders/14QPe3o6LeSjYi-kGhCZJM4pn-I55YsJ?usp=sharing>. Tugas **clustering** (*unsupervised Learning*) adalah mengelompokkan pelanggan berdasarkan data pelanggan di dealer tanpa memperhatikan label kelas apakah pelanggan tertarik untuk membeli kendaraan baru atau tidak. Tugas **classification** (*supervised learning*) adalah memprediksi apakah pelanggan tertarik untuk membeli kendaraan baru atau tidak berdasarkan data pelanggan di dealer.
2. Untuk setiap *task*, mahasiswa **diharuskan** melakukan enam hal berikut:
 - a. Formulasi Masalah: jelaskan apa permasalahan yang akan diselesaikan.
 - b. Eksplorasi dan Persiapan Data (termasuk *data splitting*): lakukan semua teknik eksplorasi dan persiapan data yang menurut Anda perlu dilakukan. Jelaskan mengapa perlu melakukan teknik tersebut, dan lakukan analisis terhadap hasilnya.
 - c. Pemodelan: bangunlah model menggunakan data hasil praproses 2.b, dan lakukan proses *training* untuk mendapatkan hasil terbaik. Jelaskan secara detail semua proses yang Anda lakukan dilengkapi dengan justifikasi dan analisis hasilnya.
 - d. Evaluasi: pilih metode evaluasi yang sesuai beserta justifikasinya. Lakukan evaluasi terhadap model yang telah dihasilkan. Berikan analisis terhadap hasil evaluasi.
 - e. Eksperimen: lakukan berbagai eksperimen yang melibatkan tahapan Eksplorasi dan Persiapan Data, Pemodelan, dan Evaluasi untuk mendapatkan hasil terbaik.

Laporkan semua Eksperimen yang Anda lakukan beserta analisis hasil dan perbandingannya.

- f. Kesimpulan: berikan kesimpulan dari semua proses yang dijalankan beserta hasil akhir dari berbagai eksperimen yang telah dilakukan.
3. Untuk tugas tahap satu (*clustering*), mahasiswa harus membangun program sendiri (*coding from scratch*) dalam pemodelan. **Tidak diperbolehkan** menggunakan library.
4. Bahasa pemrograman yang digunakan adalah **Python**. **Tidak diperbolehkan** menggunakan bahasa pemrograman lainnya.
5. Pengumpulan tugas berupa:
 - a. Laporan diketik dalam format file .pdf pada kertas berukuran A4 dengan margin 3 cm, spasi 1.15, menggunakan Times New Roman Size 12.
 - b. *Source code* yang siap dieksekusi (.py/.ipynb)
 - c. Video presentasi dengan durasi 10 menit yang diunggah pada akun youtube. Sertakan tautan (link) pada laporan pada poin (a)
 - d. Semua Dataset yang dihasilkan dari setiap tahapan eksplorasi harus dicantumkan pada laporan poin (a)

Penilaian:

1. Laporan (30%):
 - a. Kelengkapan bagian laporan (5%)
 - b. Kesesuaian laporan dengan hasil hasil luaran *source code* (5%)
 - c. Analisis pemilihan Teknik yang digunakan beserta interpretasi hasil (15%)
 - d. Akurasi prediksi (5%)
2. *Source code* (40%):
 - a. Ketepatan solusi beserta penjelasannya. *Source code* wajib diberikan komentar sebaik-baiknya supaya mudah dipahami oleh dosen dan asisten dosen.
3. Presentasi (30%):
 - a. Organisasi konten presentasi (5%)
 - b. Pemahaman tentang projek (25%)