**Deskripsi Tugas Explorasi AutoML**

**MK Pembelajaran Mesin Lanjut**

**Ketentuan Umum:**

1. Tugas dikerjakan secara berkelompok **3-4 mahasiswa**, di mana **pembagian kelompok bebas**.
2. Setap kelompok harus **memillih satu datase**t dari list di bawah. Tidak boleh ada kelompok yang menggunakan dataset yang sama.
3. Harus mematuhi etika plagiasi baik ide, kode, maupun laporan.
4. Deadline pengumpulan tugas adalah **Jumat, 31 Mei 2024, pukul 23.59 WIB.**

**Deskripsi Tugas:**

1. Dataset yang digunakan dalam tugas besar adalah salah satu dari dataset:
   1. Ecoli <https://www.openml.org/search?type=data&sort=runs&id=39&status=active>
   2. heart-h <https://www.openml.org/search?type=data&status=active&id=963>
   3. vote <https://www.openml.org/search?type=data&status=active&id=56&sort=runs>
   4. wine <https://www.openml.org/search?type=data&status=active&id=43571>
   5. vehicle <https://www.openml.org/search?type=data&status=active&id=994>
   6. pima <https://www.openml.org/search?type=data&status=active&id=43582>
   7. tic-tac-toe <https://www.openml.org/search?type=data&status=active&id=50>
   8. cars <https://www.openml.org/search?type=data&status=active&id=455>
   9. breast-w <https://www.openml.org/search?type=data&status=active&id=15>
   10. eucalyptus <https://www.openml.org/search?type=data&status=active&id=188>
   11. glass <https://www.openml.org/search?type=data&status=active&id=41&sort=runs>
2. Setiap kelompok harus memilih **dataset yang berbeda**
3. Setiap kelompok harus menyelesaikan sebuah task terhadap dataset tersebut dengan menerapkan Automated Machine Learning.
4. Machine Learning Automation Tools yang harus digunakan adalah
   1. **TPOT** <http://epistasislab.github.io/tpot/>, dan
   2. **Niapy** <https://niapy.org/en/stable/>
5. Bahasa pemrograman yang digunakan adalah Python.
6. Untuk task tersebut, mahasiswa diharapkan setidaknya melakukan hal-hal berikut:
   1. Formulasi Masalah: mengacu pada 4 tahap From Problem to ML Solution.
   2. Eksplorasi dan Persiapan Data (termasuk split data): Setidaknya menggunakan dua buah teknik eksplorasi dan/atau penyiapan data. Jelaskan mengapa menggunakan teknik tersebut.
   3. Pemodelan: Jelaskan Langkah detail kedua Machine Learning Automation Tools tersebut.
   4. Eksperimen: Wajib melakukan setidaknya sebuah eksperimen dengan menggunakan kedua tools yang disebutkan.
   5. Evaluasi: Pilih metode evaluasi yang ingin digunakan serta jelaskan alasannya. Lakukan evaluasi terhadap model yang digunakan. Berikan analisis terhadap hasil evaluasi.
   6. Bandingkan Langkah dan hasil kedua tool tersebut.
   7. Kesimpulan: Sampaikan kesimpulan dari semua proses yang dijalankan, apa saja lesson learned, dan apa saran untuk improvement ke depan.
7. Delivery tugas berupa:
   1. Laporan diketik dalam A4 menggunakan times new roman size 12, margin 3cm, space 1.15 (.pdf)
   2. Kode-kode yang siap dieksekusi (.py/.ipynb)
   3. Dataset hasil eksplorasi

**Penilaian:**

1. Laporan (50%):
   1. Sesuai Format (5%)
   2. Penggunaan Bahasa yang baik dan benar (10%)
   3. Kelengkapan (25%)
   4. Akurasi laporan terhadap yang dikerjakan/kode (10%)
2. Kode (50%):
   1. Kelengkapan semua tahapan (30%)
   2. Eksekusi kode (10%)
   3. Kerapian kode, penggunaan fungsi secara tepat, pemberian komentar (10%)
3. Bonus untuk extraordinary results (20%):
   1. Melakukan lebih banyak eksperimen yang menarik. Pembuatan kode dengan dokumentasi yang sangat lengkap.
   2. Penyajian laporan yang menarik, penyertaan ilustrasi yang membantu pemahaman pembaca, analisis yang tajam, usulan yang bermanfaat.