

# Tugas Besar IF3170

## Inteligenzia Buatan 2017/2018

### Aplikasi Web Prediksi Income Per Tahun

#### Pendahuluan

Tugas besar ini dikerjakan secara berkelompok dengan anggota 4-5 orang per kelompok (boleh lintas kelas). Silakan isi anggota kelompok pada form ini sebelum **Sabtu 11 November 2017 pukul 23.59**: [https://docs.google.com/a/std.stei.itb.ac.id/spreadsheets/d/1w1n3WdeOxTYkiou0vo0R2jiysHk5Kyx6\\_alpJLHUZDs/edit?usp=sharing](https://docs.google.com/a/std.stei.itb.ac.id/spreadsheets/d/1w1n3WdeOxTYkiou0vo0R2jiysHk5Kyx6_alpJLHUZDs/edit?usp=sharing).

#### Deskripsi Persoalan

Mahasiswa diminta untuk membuat sebuah aplikasi web yang dapat memberikan prediksi income per tahun berbasis pembelajaran mesin. Aplikasi yang dibuat harus dapat menerima masukan berupa semua atribut pada dataset. Silahkan baca CencusIncome.names.txt untuk informasi detilnya. Selanjutnya, sistem akan memberikan hasil prediksi berupa  $>50K$  atau  $\leq 50K$ . Pelaksanaan tugas ini akan dilakukan dalam 2 tahap yaitu eksperimen dan pengembangan aplikasi web.

##### A. Eksperimen untuk mendapatkan model terbaik

Lakukanlah analisis data, desain skenario eksperimen, dan eksperimen untuk menentukan konfigurasi agar didapat model terbaik (dengan semua algoritma pembelajaran yang sudah pernah diberikan di kuliah IF3170), dari dataset CencusIncome dengan menggunakan skema 10-fold cross validation.

Buatlah laporan per kelompok yang menjelaskan hal-hal berikut ini:

1. Hasil analisis data yang dilakukan dan penanganan apa saja yang harus dilakukan. Analisis termasuk menentukan ukuran kinerja yang akan digunakan dalam eksperimen. Jika tidak ada penanganan khusus, tuliskan secara eksplisit.
2. Lakukanlah eksperimen yang dilakukan untuk mendapatkan algoritma pembelajaran terbaik dengan menggunakan 10-fold cross validation untuk CencusIncome.data.txt. Lakukanlah analisis akurasi dan confusion matrix terhadap model terbaik yang didapatkan.
3. Lakukanlah full training dengan CencusIncome.data.txt menggunakan algoritma pembelajaran terbaik, lalu simpanlah model yang dihasilkan.
4. Load model yang telah disimpan dari langkah 3, dan lakukanlah evaluasi model tersebut dengan memprediksi data pada CencusIncome.test.txt.

Tugas tahap A dikumpulkan ke asisten saat asistensi berupa hasil download notebook dalam dua format yaitu file .ipynb dan pdf. Penamaan file yang dikumpulkan: Tubes2A\_[NIM terkecil anggota].zip (misal: Tubes2A\_13515001.zip yang berisi Tubes2A\_13515001.ipynb dan Tubes2A\_13515001.pdf).

B. Pembangunan aplikasi web yang memanfaatkan model terbaik dari tahap A

Aplikasi web harus dapat menerima masukan berupa nilai semua atribut, dan memberikan hasil prediksi berupa  $>50K$  atau  $\leq 50K$ . Aplikasi web ini harus memanfaatkan model terbaik yang didapatkan dari tahap A.

## **Pengumpulan dan Demo Program**

Tahap A (termasuk asistensi) dilakukan paling lambat tanggal 20 November 2017 (pengumpulan ke server paling lambat jam 06.55 pada tanggal tersebut).

Proses pengumpulan dan demo program dilakukan 3 minggu setelah pengumuman tugas (tugas diberikan terhitung tanggal 8 November 2017). Dengan demikian, pengumpulan dilakukan pada tanggal 28 November 2017 melalui situs kuliah <https://stei.kuliah.itb.ac.id/course/view.php?id=36>. Lebih detail mengenai jadwal asistensi maupun demo akan diatur oleh asisten.

## **Catatan Tambahan**

Selama pelaksanaan tugas ini, seluruh mahasiswa dimohon untuk memperhatikan hal-hal berikut.

1. Jika ada hal yang ingin ditanyakan mengenai spek tugas atau hal-hal lain yang berhubungan dengan tugas, mohon untuk mengajukan pertanyaan pada milis dan tidak bertanya secara langsung kepada asisten. Asisten tidak akan melayani pertanyaan di luar milis untuk menghindari penyebaran informasi yang tidak merata.
2. Bagi mahasiswa yang bukan merupakan angkatan 2015, dimohon untuk bergabung ke dalam 1 kelompok yang sama (meskipun jumlah anggotanya dapat melebihi 5 orang).