**Eksperimen**

**IF4071 Pembelajaran Mesin**

**Tanggal: 30 November 2017**

**Aturan:**

1. Eksperimen menggunakan Decision Tree Learning & Artificial Neural Network
2. Eksperimen dilakukan secara berkelompok 3 orang.
   1. *Dilarang bekerja sama dengan kelompok lain.*
   2. Asisten dapat meminta kelompok yang bekerja sama untuk keluar dari ruangan dan tidak perlu mengumpulkan hasil eksperimen.
   3. *Anggota kelompok yang tidak hadir tidak boleh dimasukkan namanya di dalam laporan*
3. Kakas yang digunakan adalah sci-kit learn untuk pembangunan model DTL dan ANN pada Jupiter notebook yang sudah dibuat scriptnya pada tugas kecil I.
4. Penggunaan modul pembelajaran kelompok lain merupakan suatu kecurangan.
5. Tidak boleh menggunakan internet selama eksperimen dilakukan.
6. Waktu eksperimen maksimum 2,5 jam (13.00-15.30).

Diberikan dataset CensusIncome. Lakukanlah analisis data, desain skenario eksperimen, dan eksperimen untuk menentukan konfigurasi agar didapat model terbaik dari kedua dataset dengan menggunakan skema 10-fold cross validation.

Buatlah laporan per kelompok yang menjelaskan hal-hal berikut ini dari setiap dataset:

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. Skenario eksperimen yang dilakukan berdasarkan hasil analisis data (mencakup faktor apa saja yang perlu dipertimbangkan untuk mendapat model terbaik), termasuk pembagian data dengan menggunakan 10-fold cross validation.
3. Hasil eksperimen sesuai skenario eksperimen yang dilakukan, dan tentukanlah konfigurasi dari dua model terbaiknya.
4. Hasil analisis fitur dengan visualisasi hubungan setiap fitur dengan kelas.
5. Hasil analisis terhadap dua model terbaik yang didapatkan untuk mengecek apakah kedua model tersebut konsisten, dan apakah model terbaik ranking 1 memenuhi sifat hipotesis yang lebih general atau sama dengan dibandingkan model terbaik ranking 2.
6. Buatlah kesimpulan dari hasil yang didapatkan.

Penilaian dilakukan berdasarkan laporan yang dibuat oleh setiap kelompok. Nilai akhir adalah rata-rata dari nilai kedua dataset.

***Selamat mengerjakan.***