**Departemen Teknik Elektro, Program Studi S2 Teknik Elektro**

**Ujian Akhir Semester**

**Semester Genap 2020/2021**

|  |  |  |  |
| --- | --- | --- | --- |
| Mata Kuliah (Kode) | : Rekayasa Data Terapan (ENEE802606) | | |
| Dosen | : Dr. Prima Dewi Purnamasari, Dr. Anak Agung Putri Ratna | | |
| Hari, Tanggal | : Jumat, 18 Juni 2021 | Waktu | : 7 hari. Dikumpulkan Jumat 25 Juni 2021 jam 09.00 WIB |
| Kompetensi (CPL) | : 5 | Sifat Ujian | : Proyek Individu |

**Petunjuk Ujian**

Kerjakan ujian ini dalam bentuk Notebook Python, dengan bagian penjelasan dituliskan dalam bentuk “markdown text” dan bagian program dituliskan dalam bentuk kode Python. Kumpulkan melalui area yang tersedia di EMAS. **File yang dikumpulkan berjumlah 3 buah:**

1) sourcecode dalam bentuk .ipynb diberi nama file: nama anda\_code.ipynb, dibuat dengan Google Colab

2) file data training dalam bentuk .csv dengan nama file: nama anda\_train.csv, dan

3) file data testing dalam bentuk .csv dengan nama file: nama anda\_test.csv.

Presentasi proyek dilaksanakan sesuai jadwal UAS dengan alokasi waktu masing-masing mahasiswa 5 menit (kelebihan waktu akan mengurangi nilai)

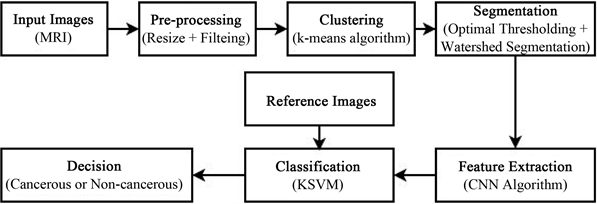
**Deskripsi umum Proyek Akhir:**

Rancanglah sebuah solusi untuk **permasalahan penilaian ujian esai otomatis** dari data yang diberikan dengan metode text processing dan machine learning. Tahapan Text Processing (pre-processing & feature extraction) dan classifier dapat didesain secara bebas namun sesuai dengan yang telah dipelajari di kelas ini dan akan menjadi penilaian utama. Output akhir yang diharapkan dari sistem adalah nilai yang mendekati nilai human rater. Anda bebas apakah akan memberi output berupa kelas (kategorikal), atau angka (mana yang menurut Anda lebih cocok).

**Ketentuan:**

1. Rancangan dituliskan dengan cara menggambarkan langkah-langkah desain dalam bentuk *block diagram* (image dapat di-*insert* dalam area *markdown text*) dan penjelasan berupa teks yang menjelaskan gambar. Langkah-langkah yang dituliskan harus lengkap sesuai dengan apa yang akan diimplimentasikan dalam program. Bagian *Pre-processing* dan *Feature Extraction* sebagian besar (atau seluruhnya) mengikuti apa yang dipelajari di kelas dan harus di-*break down* sesuai dengan langkah yang akan dilakukan (bukan seperti gambar di bawah).

*Ket: block diagram bukanlah flowchart, melainkan hanya kumpulan blok yang menunjukkan urutan kerja. Contoh block diagram adalah sebagai berikut:*



1. Bagilah data yang diberikan ke dalam 2 file CSV terpisah, yaitu untuk data training, dan untuk data testing. Data training dan data testing dibagi dengan proporsi 80% untuk data training, dan 20% untuk data testing. Proses pemisahan data ini dapat dilakukan dengan software apapun, tidak harus Python, misal menggunakan Matlab, Ms Excel dsb. Tuliskan proses yang Anda lakukan dalam bentuk markdown text. Jelaskan atribut dan kelas data. File data training, dan data testing dilampirkan Ketika pengumpulan ujian.
2. Implementasikan rancangan pada nomer 1 dalam kode Python dengan penjelasan yang memadai di tiap blok prosesnya, dalam bentuk markdown text. Jadi, kode program dipisah per blok proses. Data pada no 2 di-*insert* ke dalam program dalam bentuk upload file, agar langsung dapat dijalankan di **Google Colab,** contoh:

# It will prompt you to select a file. Click on “Choose Files” then select and upload the file. # Wait for the file to be 100% uploaded.

from google.colab import files

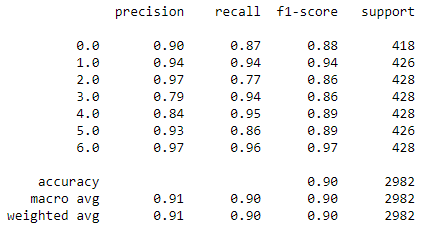
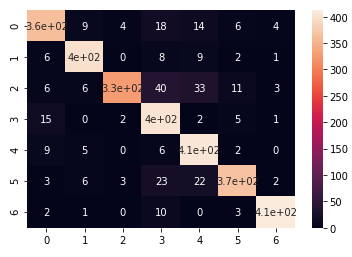
uploaded = files.upload()

# To store dataset in a Pandas Dataframe

import io

df2 = pd.read\_csv(io.BytesIO(uploaded['Filename.csv'])

1. Tampilkan hasil klasifikasinya dalam bentuk tabel dan confusion matrix, serta grafik lainnya jika dirasa perlu, contoh:

1. Buatlah analisis dan kesimpulan dari hasil pada nomer 4. Anda dapat melakukan analisis pada hasil mana akurasi rendah dan kira-kira apa penyebabnya.
2. Jika Anda menyadur sebagian atau seluruh langkah yang digunakan untuk mengerjakan ujian ini dari Internet, maka tuliskan referensinya dengan lengkap, dengan sitasi yang memadai (gunakan IEEE Style). Misal untuk pengolahan data dari referensi A, untuk klasifikasi dari referensi B dsb.