

Tugas Pekan ke-10 *Text Classification and Sentiment Analysis*

Batas pengumpulan: Jumat 27 November 2020, pukul 10.59 pagi, melalui LMS

Deskripsi

Lakukan implementasi dari salah satu metode klasifikasi teks (*supervised*) berikut (pilih salah satu):

- Support vector machines (SVM)
- Naïve Bayes
- k-NN
- Decision tree

Gunakan dataset berbahasa **Indonesia** dengan tema bebas. Cantumkan dan sertakan sumber datanya.

Program dan Analisis yang Harus Dibuat

1. Buatlah program untuk mengklasifikasikan teks menggunakan salah satu algoritma *supervised* di atas.
2. Lakukan evaluasi dan analisis hasil klasifikasi berdasarkan Precision, Recall, dan F-Measure.
3. Lakukan analisis terhadap kinerja algoritma klasifikasi berdasarkan:
 - a. Pengaruh persentase data latih dan data uji
 - b. Waktu pelatihan yang diperlukan
 - c. Data yang salah diklasifikasikan, jelaskan mengapa data tersebut salah klasifikasi! Lampirkan juga data yang salah diklasifikasikan.

File yang harus dikumpulkan:

1. Program, berisi pendefinisian algoritma, setting parameter, dan dataset yang digunakan.
2. Laporan: 1 file pdf, maksimum panjang laporan adalah 4 halaman. Laporan berisi jawaban pertanyaan nomor 2 yang menjelaskan hasil evaluasi dan analisis hasil klasifikasi berdasarkan Precision, Recall, dan F-Measure, serta jawaban pertanyaan nomor 3.

Detail penilaian:

- a. Program:
 - Kebenaran implementasi algoritma klasifikasi teks menggunakan salah satu algoritma yang telah ditentukan di atas [30 poin]
- b. Laporan:
 - kelengkapan jawaban nomor 2 dan 3 [20 poin]
 - analisis jawaban nomor 2 [20 poin]
 - analisis jawaban nomor 3 [30 poin]

Jika ada pertanyaan, silakan disampaikan melalui *channel* pekan_10_text_classification di slack.