

Mata Kuliah : IBDA3111 / Prapemrosesan dan Rekayasa Data
Hari, Tanggal : XX, XX
Waktu : XX
Sifat Ujian : Proyek Individu

PETUNJUK

- Kerjakanlah soal-soal di bawah ini lalu dikumpulkan dalam format text entry, .ipynb, dan .pdf dengan mengikuti rubrik yang ada lalu dikumpulkan di dalam Canvas bagian *assignments*, bagian UTS.
- Setelah selesai mengerjakan ujian, salinlah pernyataan komitmen integritas berikut ini di dalam file laporan (.pdf) dan bubuhkanlah tanda tangan di bawahnya:

"Di hadapan TUHAN yang hidup, saya menegaskan bahwa saya tidak memberikan maupun menerima bantuan apapun—baik lisan, tulisan, maupun elektronik—di dalam ujian ini selain daripada apa yang telah diizinkan oleh pengajar, dan tidak akan menyebarkan baik soal maupun jawaban ujian kepada pihak lain."

PENGANTAR

Terdapat banyak sekali data di sekitar kita yang dapat kita manfaatkan untuk memberkati sesama kita. Oleh karena itu peserta kelas IBDA3111 diminta untuk berpartisipasi aktif dengan menggunakan ilmu yang sudah diperoleh di kelas ini untuk mengidentifikasi ketersediaan data, melakukan pemrosesan dan pengolahan data, serta menggunakannya untuk menawarkan solusi/analisa bagi permasalahan tertentu. Proyek ini akan menjadi proyek UTS kalian dimana proyek ini akan dikerjakan di dalam 3 tahap, yaitu proposal singkat mengenai pilihan data, pemrograman, dan laporan. Kumpulkanlah semua bagian sebelum tenggat waktu yang diberikan.

DESKRIPSI TUGAS

1. Proposal Data (20 poin)

Sebelum melakukan pengolahan data, diperlukan pertimbangan yang matang tentang data apa yang dipilih dan solusi potensial apa yang dapat ditawarkan dengan menggunakan data pilihan kalian. Maka 1 minggu sebelum proyek ini dimulai:

- Kumpulkanlah sebuah proposal singkat (kurang dari 500 kata),
- Mengenai data yang akan anda gunakan untuk proyek UTS,
- Melalui text entry yang sudah disiapkan di canvas.

Proposal akan dinilai secara kasar terlebih dahulu untuk memastikan semua data unik:

- Jika terdapat pilihan data yang sama, hanya proposal dari mahasiswa yang mengumpulkan terlebih dahulu yang diterima.
- Jika proposal lengkap dan valid, maka otomatis akan mendapatkan nilai 20 poin.
- Bonus 10 poin yang akan diberikan kepada mahasiswa yang memperoleh data secara tidak langsung (crawling/scraping)

2. Tugas pemrograman (40 poin)

- Dengan menggunakan Jupyter Notebook lakukanlah proses pengolahan data untuk sektor yang Anda pilih
- dengan ketentuan perlu menerapkan **minimal 2 metode** dari *data cleaning* dan **2 metode** dari *feature selection* yang sudah diajarkan di kelas.
- Setelah diproses, anda bebas memilih salah satu dari 2 pendekatan berikut terhadap data tersebut untuk menyelesaikan permasalahan yang dipilih:
 - analisis (statistik dan visualisasi)
 - pemodelan (*machine learning*)
- Anda juga diminta menjelaskan secara lebih detail pekerjaan yang Anda lakukan dengan memanfaatkan fitur markdown dari Jupyter Notebook supaya diperoleh tampilan yang lebih mudah dibaca.

3. Tugas laporan (40 poin)

Anda perlu membuat laporan dalam bentuk pdf dengan ketentuan sebagai berikut:

- Judul
- Latar belakang: paparan singkat studi kasus permasalahan yang anda pilih
- Data: analisis kebutuhan data (data yang digunakan dan teknik prapemrosesan & rekayasa data yang dilakukan), daftar dan penjelasan dari open data & data eksternal yang digunakan (jika ada)
- Solusi: hasil dan penjelasan dari solusi yang ditawarkan. Solusi dapat berbentuk analisa dan visualisasi data (menceritakan hasil analisa data Anda) atau pemodelan (*machine learning*)
- Kesimpulan

Gunakan template laporan standard dari Calvin Institute of Technology

RUBRIK PENILAIAN

Proposal: masalah yang coba diselesaikan	0 poin	5 poin
	Tidak ada	Ada dan valid
Proposal: deskripsi singkat data	0 poin	5 poin
	Tidak ada	Ada dan valid
Proposal: sumber data	0 poin	5 poin
	Tidak ada	Ada dan valid
Proposal: cara memperoleh data	0 poin	5 poin
	Tidak ada	Ada dan valid

	0 poin	5 poin	10 poin
Pemrograman: kemampuan data cleaning	Program yang dibuat tidak menunjukkan metode data cleaning	Program yang dibuat menunjukkan 1 metode data cleaning	Program yang dibuat menunjukkan 2 metode data cleaning
Pemrograman: kemampuan feature selection	0 poin Program yang dibuat tidak menunjukkan metode feature selection	10 poin Program yang dibuat menunjukkan 1 metode feature selection	10 poin Program yang dibuat menunjukkan 2 metode feature selection
Pemrograman: solusi yang ditawarkan berhasil dieksekusi	0 poin Program tidak dapat dieksekusi	5 poin Program yang dikumpulkan sudah dieksekusi tetapi ada beberapa bagian yang error	10 poin Program yang dikumpulkan berhasil dieksekusi
Pemrograman: Kerapihan penjelasan di dalam bagian <i>markdown</i>	0 poin Tidak ada penjelasan di bagian <i>markdown</i>	5 poin Penjelasan di bagian <i>markdown</i> tidak jelas dan tidak rapi.	10 poin Penjelasan di bagian <i>markdown</i> jelas dan rapi

	0 poin	5 poin	10 poin	20 poin
Laporan: Kemampuan komunikasi	Tidak ditemukan inti gagasan apapun dari laporan.	Gagasan utama yang disampaikan kurang jelas atau sulit dipahami.	Gagasan utama yang disampaikan jelas dan mudah dipahami.	Gagasan utama yang disampaikan jelas dan mudah dipahami disertai dengan sistematika penyampaian yang kuat.
Laporan: Konten	0 poin Tidak ditemukan pembahasan apapun mengenai topik proyek	5 poin Data dan analisa/modal yang digunakan tidak sesuai dengan permasalahan yang ingin diselesaikan	10 poin Data dan analisa/model yang digunakan dapat digunakan untuk permasalahan yang ingin diselesaikan, namun solusinya kurang komprehensif	20 poin Data dan analisa/model yang digunakan sesuai untuk permasalahan yang ingin diselesaikan, dan solusinya komprehensif