**Projek Akhir MK Data Mining – Miniriset**

**Mata Kuliah: Data Mining**  
**Topik:** Scrap Data, Text Mining, Analisis Sentimen, dan Penyusunan Artikel

1. **Tujuan Tugas**

Mahasiswa diharapkan dapat menguasai teknik pengumpulan data (scrap data) dari sumber online, menerapkan metode text mining untuk mengolah data teks, melakukan analisis sentimen untuk memahami opini publik, dan menyusun artikel ilmiah berdasarkan hasil riset.

1. **Deskripsi Tugas**
2. **Data Collection:** Scrap data teks dari sumber online (misalnya Twitter, marketplace, playstore, maps, review hotel, review produk, dsb), jumlah data minimal 500.
3. **Text Mining:** Lakukan preprocessing teks (cleaning, tokenizing, stopword removal, stemming/lemmatization, dll sesuai kebutuhan).
4. **Sentiment Analysis:** Terapkan metode analisis sentimen untuk mengkategorikan data (positif, negatif, atau netral) dengan metode-metode data mining. Lakukan analisis pada sentiment yang didapatkan, most word, most negative/positif, dll.
5. **Artikel:** Susun artikel ilmiah dengan struktur lengkap: Judul, Identitas, Abstrak, Latar Belakang, Metode, Hasil & Pembahasan, Kesimpulan, dan Daftar Pustaka.
6. **Panduan Penulisan Artikel (Target Sinta 5,6)**

| Judul | Judul Memuat Metode dan Data yang dibahas |
| --- | --- |
| Penulis | Mahasiswa, Dosen, Asisten  Dosen: Budi Prasetiyo ([bprasetiyo@mail.unnes.ac.id](mailto:bprasetiyo@mail.unnes.ac.id))  Asisten: Regina Ayumi Ulayyaa ([reginaayumi@students.unnes.ac.id](mailto:reginaayumi@students.unnes.ac.id)) |
| Abstrak |  |
| Latar Belakang (setiap paragraph disertai sitasi) | 1. Paragraf 1: Cerita mulai masalah umum, ke khusus (jangan terlalu panjang lebar). 2. Paragraf 2: Cerita masalah apa yg akan saudara angkat, disertai metode yg sdh ada/solusi yg sdh ada/penelitian terkait. 3. Paragraf 3: Cerita penelitian sebelumnya yang mendekati/mirip2, metode2 yg digunakan apa, keunggulan metode tsb. Sehingga metode ini yg akan saudara gunakan. Tarik ke masalah saudara yg akan di selesaikan sesuai judul yg diangkat dengan metode tsb 4. Pargaraf 4: Gap Riset: Uraikan perbedaan dgn penelitian sebelumnya. Misal peneliti sebelumnya sdh ada yg membahas A,B,C tetapi belum ada yg membahas D. Maka saudara yg akan membahas D ini. Peneliti 1 membahas A, peneliti 2 membahas B, peneliti 3 membahas C. Namun belum ada yg membahas AB, atau BC, atau belum ada yg bahas D. Maka saudara akan melakukan D ini. 5. Paragraf 5 (Penutup): Paragraf terakhir ceritakan apa yg akan dilakukan. Mulai dr input, proses, output, datanya apa, jml datanya brp, yg akan diukur apa, pakai tools apa. Intinya memberitahu pembaca apa yang akan dilakukan. |
| Metode | 1. Buat bagan/flowchart apa yang akan saudara lakukan. 2. Metode (Ungkapkan apa yg saudara lakukan, mulai input-proses-output, tool yg digunakan) |
| Hasil & pembahasan | 1. Tampilkan hasil, outout table/chart/grafik. Bahas grafik tersebut 2. Untuk kasus analisis sentiment focus analisis mendalam pada temuan topik2 nya, bisa juga dibandingkan hasil antara n-gram, n=1, n=2 |
| Simpulan | Simpulan hasil riset |
| Daftar Pustaka | 1. Minimal 5 Jurnal Internasional 2. Total sitasi 15-20 (minimal 15) |

Artikel Rujukan Utama:

**“Analisis Sentimen pada review Aplikasi Grab di Google Play Store Menggunakan Support Vector Machine”**

[Unduh](https://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/9681) disini: <https://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/9681/pdf>

1. **Pengumpulan**

Mengisi lewat Sheet di Folder Drive

1. **Rubrik Penilaian (menyusul)**

| Judul | Judul Memuat Metode dan Data/objek yang dibahas | terpenuhi: 10, tidak: 0 |
| --- | --- | --- |
| Penulis | Mahasiswa, Dosen, Asisten  Dosen: Budi Prasetiyo ([bprasetiyo@mail.unnes.ac.id](mailto:bprasetiyo@mail.unnes.ac.id))  Asisten: Regina Ayumi Ulayyaa ([reginaayumi@students.unnes.ac.id](mailto:reginaayumi@students.unnes.ac.id)) |  |
| Abstrak | Abstrak memuat: permasalahan singkat, data, metode, tool, hasil | lengkap: 10, kurang lengkap: 5, tidak ada: 0 |
| Latar Belakang (setiap paragraph disertai sitasi) | 1. Paragraf 1: Cerita mulai masalah umum, ke khusus (jangan terlalu panjang lebar). 2. Paragraf 2: Cerita masalah apa yg akan saudara angkat, disertai metode yg sdh ada/solusi yg sdh ada/penelitian terkait. 3. Paragraf 3: Cerita penelitian sebelumnya yang mendekati/mirip2, metode2 yg digunakan apa, keunggulan metode tsb. Sehingga metode ini yg akan saudara gunakan. Tarik ke masalah saudara yg akan di selesaikan sesuai judul yg diangkat dengan metode tsb 4. Pargaraf 4: Gap Riset: Uraikan perbedaan dgn penelitian sebelumnya. Misal peneliti sebelumnya sdh ada yg membahas A,B,C tetapi belum ada yg membahas D. Maka saudara yg akan membahas D ini. Peneliti 1 membahas A, peneliti 2 membahas B, peneliti 3 membahas C. Namun belum ada yg membahas AB, atau BC, atau belum ada yg bahas D. Maka saudara akan melakukan D ini. 5. Paragraf 5 (Penutup): Paragraf terakhir ceritakan apa yg akan dilakukan. Mulai dr input, proses, output, datanya apa, jml datanya brp, yg akan diukur apa, pakai tools apa. Intinya memberitahu pembaca apa yang akan dilakukan. | * Susunan Latar Belakang sesuai alur: 10, kurang sesuai 5 * Semua paragraf ada sitasi: 10, tidak ada: 0 * Ada paragraf yang Menceritakan keunggulan metode yg dipilih. ada: 10, tidak ada: 0 * Paragraf penutup: detail:10, cukup: 5, tidak ada penutup: 0 |
| Metode | 1. Buat bagan/flowchart apa yang akan saudara lakukan. 2. Metode (Ungkapkan apa yg saudara lakukan, mulai input-proses-output, tool yg digunakan) | * Terdapat bagan flowchart, ada; 10, tidak ada: 0 * Metode detail, mulai dari: data collection, pre-procesing, data anaisis, (lengkap: 10, cukup: 5) * Menyajikan sampel contoh data before-after (misal sebelum preprocessing vs setelah, sample data mentah vs setelah di cleaning). Terpenuhi: 5, tidak: 0 * Ada sample contoh potongan sourcecode. Ada: 5, tidak ad:0 |
| Hasil & pembahasan | 1. Tampilkan hasil, outout table/chart/grafik. Bahas grafik tersebut 2. Untuk kasus analisis sentiment focus analisis mendalam pada temuan topik2 nya, bisa juga dibandingkan hasil antara n-gram, n=1, n=2 | * Ada wordcloud + dibahas + diberi caption + disebut diparagraf. Semua terpenuhi: 10, bberap terpenuhi: 5, tidak ada wordcloud: 0 * Grafik/Chart + dibahas + diberi caption + disebut di paragraf. Lengkap: 10, kurang lengkap: 5, tidak ada chart: 0 * Tabel |
| Simpulan | Simpulan hasil riset |  |
| Daftar Pustaka | 1. Minimal 5 Jurnal Internasional 2. Total sitasi 15-20 (minimal 15) | * Minimal >= 5 jurnal internasional: 10, 2-4: 5; tidak ada: 0 * Total sitasi minimal 15: skor 10 * Total sitasi 0-5: skor 0 * Total sitasi 5-10: skor 5 * Sitasi 11-15: skor 7,5 |