

**MODUL PRAKTIKUM  
BUSINESS INTELLIGENCE**

**MODUL KE : \_\_\_\_\_**

**NAMA : \_\_\_\_\_**

**NPM : \_\_\_\_\_**

**KELAS: \_\_\_\_\_**

**INSTITUT INFORMATIKA DAN BISNIS DARMAJAYA**

**2025**

## MODUL X

### Text Mining & Sentiment Analysis untuk Keputusan Bisnis

#### PETUNJUK UMUM

1. Ujian ini bersifat praktikum individu.
2. Mahasiswa wajib:
  - Menyusun dataset sendiri
  - Menjalankan analisis
  - Menarik kesimpulan bisnis
3. Tidak ada satu jawaban benar mutlak, penilaian berdasarkan logika dan konsistensi.
4. Gunakan Google Colab untuk coding Python.
5. Dataset harus disimpan dalam format CSV.
6. **Dilarang: Copy hasil teman / dataset identik**

#### STUDI KASUS

Sebuah perusahaan e-commerce ingin mengetahui sentimen pelanggan terhadap salah satu produk berdasarkan review pelanggan pada periode Oktober 2025. Hasil analisis akan digunakan untuk menentukan strategi promosi dan perbaikan produk.

#### TUGAS UAS

##### BAGIAN A – PENYUSUNAN DATA (30 POIN)

1. Pilih 1 produk (bebas tapi tidak boleh identic antar mahasiswa)
2. Buat dataset review produk dengan ketentuan:
  - Sumber: website/market place, sosial media masing-masing produk.
  - Periode: 1–31 Oktober 2025
  - Minimal 40 review
  - Review berupa kalimat opini
3. Susun data menggunakan Excel atau Google Sheets dengan struktur:

Tanggal	Review	Rating
---------	--------	--------

4. Dataset harus memenuhi:
  - Minimal 5 review negatif

- Minimal 5 review ambigu
5. Simpan dataset sebagai file CSV.
  6. Upload file CSV ke Google Colab.

### **BAGIAN B – PREPROCESSING & LABELING (20 POIN)**

1. Lakukan pembersihan teks:
  - Ubah huruf menjadi lowercase
  - Hilangkan tanda baca dan angka
2. Buat label sentimen dengan ketentuan:
  - 1 = Positif
  - 0 = Negatif
3. Penentuan label boleh berdasarkan:
  - Kata kunci
  - Rating
  - Kombinasi keduanya
4. Jelaskan secara singkat logika pelabelan yang Anda gunakan.

### **BAGIAN C – MODELLING & EVALUASI (25 POIN)**

1. Ubah teks menjadi data numerik menggunakan TF-IDF.
2. Bagi data menjadi (data split):
  - Data latih (training) sebanyak 70%
  - Data uji (testing) sebanyak 30%
3. Bangun model klasifikasi:
  - Logistic Regression (untuk mahasiswa dengan digit NPM terakhir: 0, 1,3,5,7)
  - Naive Bayes (untuk mahasiswa dengan digit NPM terakhir: 2,4,6, 8 dan 9)
4. Tampilkan hasil evaluasi masing-masing model:
  - Confusion Matrix
  - Accuracy

- Precision
- Recall

**Catatan: Nilai tidak ditentukan oleh akurasi tertinggi, tetapi oleh proses dan analisis.**

### **BAGIAN D – ANALISIS BISNIS (25 POIN)**

Jawab pertanyaan berikut berdasarkan hasil Anda sendiri:

1. Dalam konteks promosi:
  - Mana yang lebih berisiko: false positive atau false negative? Jelaskan berdasarkan hasil kasus.
2. Jika perusahaan hanya mengirim promo ke pelanggan dengan sentimen positif:
  - Apa dampak terhadap biaya dan pendapatan?
3. Berdasarkan hasil sentiment analysis:
  - Strategi bisnis apa yang sebaiknya dilakukan perusahaan?

### **OUTPUT YANG DIKUMPULKAN**

1. File CSV dataset
2. File Google Colab Notebook (.ipynb)
3. Ringkasan analisis bisnis (±1 halaman )

### **KRITERIA PENILAIAN (RINGKAS)**

<b>Komponen</b>	<b>Bobot</b>
<b>Dataset &amp; validitas data</b>	<b>30%</b>
<b>Preprocessing &amp; labeling</b>	<b>20%</b>
<b>Modelling &amp; evaluasi</b>	<b>25%</b>
<b>Analisis bisnis</b>	<b>25%</b>
<b>Total</b>	<b>100%</b>