

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Di era perkembangan teknologi informasi yang semakin pesat saat ini, media sosial telah menjadi platform utama bagi masyarakat untuk menyampaikan berbagai opini, kritik, dan apresiasi terhadap berbagai hal, termasuk layanan institusi pendidikan. Perguruan tinggi sebagai salah satu pilar pendidikan formal, dituntut untuk terus berinovasi dalam meningkatkan mutu pelayanan akademik maupun non-akademik guna menjaga reputasi di tengah masyarakat, mahasiswa, alumni, dan calon mahasiswa. Citra sebuah universitas saat ini tidak hanya dibentuk oleh pelayanan yang diberikan secara langsung, tetapi juga dipengaruhi oleh opini-opini yang berkembang di media sosial, yang kini menjadi salah satu indikator penting dalam menilai kualitas suatu institusi pendidikan.

Universitas Dharma Persada sebagai salah satu perguruan tinggi swasta di Jakarta, juga menjadi objek perhatian masyarakat di berbagai media sosial seperti Twitter, Instagram, dan forum daring. Opini, komentar, hingga ulasan dari mahasiswa dan masyarakat umum mengenai fasilitas, pelayanan administrasi, kebijakan akademik, dan program-program kampus menjadi masukan yang sangat berharga bagi pihak universitas. Informasi tersebut, jika dikelola dengan baik, dapat menjadi dasar dalam menyusun strategi pengembangan dan peningkatan layanan di lingkungan kampus.

Menurut Perhatian (2024), media sosial kini menjadi ruang diskusi aktif yang mampu merepresentasikan opini masyarakat terhadap berbagai isu sosial, termasuk layanan institusi. Melalui analisis sentimen media sosial, institusi dapat

mengetahui bagaimana persepsi publik secara real-time, yang sangat penting untuk bahan evaluasi layanan di berbagai sektor, termasuk pendidikan.

Analisis sentimen merupakan studi komputasi yang digunakan untuk mengenali, mengekspresikan, dan mengklasifikasikan opini, sentimen, atau pandangan yang terkandung dalam suatu teks, baik berupa positif, negatif, maupun netral. Dalam konteks layanan pendidikan, metode ini bermanfaat untuk mengidentifikasi tanggapan masyarakat terhadap suatu layanan atau kebijakan, sehingga institusi dapat memahami persepsi pengguna secara sistematis (YustantiWiyli & Atmaja Kusuma Panca Wahyu Rizqi Mas Raden, 2021).

Penerapan metode analisis sentimen di lingkungan pendidikan, khususnya di Universitas Dharma Persada, masih tergolong terbatas. Padahal, terdapat banyak opini mahasiswa yang tersebar di media sosial dan situs resmi kampus yang belum dimanfaatkan secara optimal sebagai bahan evaluasi institusi. Saat ini, proses evaluasi terhadap layanan kampus masih mengandalkan kuesioner atau survei manual yang hanya dilakukan pada periode tertentu, sehingga prosesnya relatif lambat, kurang efisien, dan berpotensi bias.

Opini-opini yang muncul di media sosial bersifat spontan, aktual, dan bisa menjadi sumber data penting untuk mengidentifikasi isu-isu yang berkembang di kalangan mahasiswa dan masyarakat luas. Oleh sebab itu, perlu dikembangkan sistem analisis sentimen berbasis kecerdasan buatan yang mampu mengolah opini tersebut secara otomatis, cepat, dan akurat untuk mendukung evaluasi layanan kampus.

Sebagai solusinya, penulis mengusulkan pengembangan sistem analisis sentimen berbasis Transformer, khususnya model IndoBERT yang terbukti efektif untuk menganalisis teks berbahasa Indonesia. IndoBERT merupakan model NLP berbasis BERT (Bidirectional Encoder Representations from Transformers) yang telah dilatih menggunakan dataset berbahasa Indonesia, sehingga mampu memahami konteks dan menginterpretasikan teks dengan lebih akurat (Wicaksono & Syah, 2024).

Penelitian ini bertujuan untuk mengklasifikasikan opini mahasiswa dan masyarakat mengenai Universitas Dharma Persada ke dalam tiga kategori sentimen, yaitu positif, negatif, dan netral. Data yang digunakan berasal dari media sosial seperti Twitter dan Instagram, serta platform internal kampus. Hasil analisis ini diharapkan dapat menjadi bahan pertimbangan dalam evaluasi mutu layanan, serta membantu pihak kampus dalam mengambil kebijakan strategis yang berbasis data aktual. Selain itu, penelitian ini juga diharapkan dapat memberikan kontribusi dalam pengembangan metode analisis sentimen berbasis AI di sektor pendidikan di Indonesia.

## **1.2 Identifikasi Masalah**

1. Banyaknya opini mahasiswa tentang perilaku Fakultas yang beredar di media sosial, namun belum terkaji secara sistematis.
2. Selama proses berlangsung, layanan Fakultas dievaluasi secara manual, lambat, dan subjektif.

### **1.3 Rumusan masalah**

Bagaimana penerapan metode Transformer berbasis IndoBERT mampu mengklasifikasikan sentimen opini publik terhadap Universitas Dharma Persada?

### **1.4 Batasan masalah**

1. Data yang digunakan dalam penelitian ini secara khusus difokuskan pada opini atau ulasan yang berasal dari media sosial dan platform digital terkait Universitas Dharma Persada. Pemilihan ulasan dari platform tersebut didasarkan pada tingginya aktivitas pengguna dalam menyampaikan pendapatnya mengenai pelayanan, fasilitas, dan kegiatan kampus, sehingga dapat merepresentasikan persepsi publik secara lebih aktual dan relevan.
2. Dalam proses analisis sentimen menggunakan metode Transformer, diperlukan sejumlah parameter atau fitur input berupa data teks opini yang telah melalui tahap preprocessing. Proses preprocessing ini mencakup case folding, cleaning, tokenizing, stopword removal, dan stemming agar data teks lebih terstruktur dan siap untuk diproses. Data opini yang digunakan dibatasi hanya pada ulasan yang berbahasa Indonesia tanpa menerjemahkan teks dari bahasa asing.
3. Metode yang diterapkan dalam penelitian ini terbatas pada algoritma Transformer dengan model IndoBERT sebagai model pretrained utama. Pemilihan IndoBERT dilakukan karena model ini telah dilatih menggunakan dataset berbahasa Indonesia, sehingga lebih sesuai untuk analisis sentimen terhadap opini masyarakat di Indonesia dibandingkan model pretrained lainnya seperti BERT multilingual atau model berbasis CNN/LSTM.

## 1.5 Tujuan Penelitian

Menganalisis dan menguji performa penerapan metode Transformer berbasis IndoBERT dalam mengklasifikasikan sentimen opini publik terhadap Universitas Dharma Persada. Penelitian ini bertujuan untuk mengetahui sejauh mana efektivitas model IndoBERT dalam memetakan sentimen masyarakat, baik positif, negatif, maupun netral, berdasarkan opini yang diperoleh dari media sosial dan platform digital.

## 1.6 Manfaat Penelitian

1. Memberikan hasil analisis sentimen opini masyarakat terhadap Universitas Dharma Persada yang lebih valid dan berbasis data, sehingga bisa dimanfaatkan sebagai dasar pengambilan keputusan dalam peningkatan kualitas pelayanan, fasilitas, dan kebijakan kampus.
2. Menyediakan pemahaman mengenai performa metode Transformer berbasis IndoBERT dalam memproses dan mengklasifikasikan teks opini berbahasa Indonesia, khususnya dalam konteks institusi pendidikan tinggi.
3. Menyajikan hasil perbandingan performa dan akurasi model IndoBERT yang dapat menjadi referensi dalam pemilihan algoritma analisis sentimen, baik untuk penelitian akademik maupun penerapan di bidang pendidikan dan layanan publik.
4. Menjadi bahan referensi ilmiah dan praktis untuk pengembangan penelitian di bidang analisis sentimen dan Natural Language Processing (NLP), khususnya yang memanfaatkan model pretrained bahasa Indonesia seperti IndoBERT dalam menganalisis opini publik

## 1.7 Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan penerapan algoritma machine learning berbasis Transformer, khususnya model IndoBERT, untuk membangun model klasifikasi sentimen opini publik terhadap Universitas Dharma Persada. Fokus metodologi ini adalah menganalisis performa IndoBERT dalam mengklasifikasikan sentimen positif, negatif, dan netral dari opini masyarakat di media sosial. Adapun tahapan metodologi dalam penelitian ini adalah sebagai berikut:

### 1. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh melalui proses scraping dari media sosial (seperti Twitter dan Instagram) dan platform digital lainnya yang memuat opini terkait Universitas Dharma Persada. Pengambilan data dilakukan dengan memanfaatkan tools scraping seperti sncrape atau crawler API yang dapat mengekstrak teks ulasan, tanggal posting, serta nama akun pengguna. Selain itu, dilakukan studi literatur terhadap jurnal-jurnal ilmiah terbaru yang relevan untuk mendukung pemilihan metode analisis sentimen dan validasi hasil penelitian.

### 2. Pra-Pemrosesan Data

Data opini yang diperoleh akan melalui serangkaian proses pembersihan dan transformasi agar siap untuk diproses oleh model machine learning. Tahapan pra-pemrosesan yang dilakukan meliputi:

- a. Case Folding, yaitu mengubah seluruh huruf dalam teks menjadi huruf kecil.

- b. Cleaning, yaitu menghapus karakter khusus, tanda baca, angka, dan emoji yang tidak diperlukan.
- c. Tokenizing, yaitu memecah teks menjadi potongan kata (token).
- d. Stopword Removal, yaitu menghapus kata-kata umum yang tidak memiliki makna penting dalam analisis sentimen.
- e. Stemming, yaitu mengubah kata ke bentuk dasarnya menggunakan library Sastrawi.

### 3. Pemisahan Dataset

Dataset yang telah diproses akan dibagi menjadi dua bagian, yaitu 80% sebagai data latih (training set) dan 20% sebagai data uji (testing set). Pembagian ini bertujuan untuk memastikan bahwa evaluasi performa model dilakukan secara objektif, dengan menguji model pada data yang belum pernah digunakan dalam proses pelatihan, sehingga dapat mengukur kemampuan generalisasi model secara adil.

### 4. Pelatihan Model IndoBERT

Penelitian ini akan menggunakan model pretrained IndoBERT dari IndoNLU/IndoBenchmark sebagai model dasar untuk analisis sentimen. Model IndoBERT dipilih karena telah dilatih khusus menggunakan dataset berbahasa Indonesia sehingga lebih efektif dalam memahami konteks kalimat dalam bahasa Indonesia. Proses pelatihan model dilakukan dengan fine-tuning pada dataset sentimen yang telah diproses. Hyperparameter seperti batch size, learning rate, dan jumlah epoch akan diatur untuk

mengoptimalkan performa model. Library Hugging Face Transformers dan PyTorch digunakan dalam tahap ini.

## 5. Evaluasi Model

Model yang telah dilatih akan dievaluasi menggunakan data pengujian.

Evaluasi dilakukan dengan menghitung beberapa metrik performa seperti:

- A. Akurasi
- B. Presisi
- C. Recall

Untuk proses evaluasi ini digunakan library scikit-learn yang menghasilkan classification report sebagai acuan tingkat akurasi dan efektivitas model dalam melakukan klasifikasi sentimen.

## 6. Deployment (Opsional)

Sebagai langkah tambahan, hasil dari model dapat diimplementasikan ke dalam aplikasi web sederhana berbasis Flask atau Streamlit, agar pengguna dapat mencoba langsung fitur klasifikasi sentimen secara realtime dengan memasukkan opini dan mendapatkan hasil klasifikasinya.

## 1.8 Sistematika Penulisan

### Bab I : Pendahuluan

Bab ini membahas latar belakang permasalahan terkait pentingnya mengetahui persepsi masyarakat terhadap Fakultas Teknik Universitas Dharma Persada melalui media sosial, khususnya Instagram. Selain itu, bab ini menjelaskan rumusan masalah, tujuan penelitian, manfaat, batasan masalah, dan sistematika penulisan agar penelitian terarah dan sesuai tujuan yang diinginkan.

### Bab II : Landasan Teori

Bab ini memuat teori-teori pendukung yang relevan dengan penelitian, seperti konsep dasar analisis sentimen, metode Transformer khususnya IndoBERT, text mining, Natural Language Processing (NLP), serta tahapan CRISP-DM. Selain itu, dimuat pula hasil penelitian terdahulu sebagai referensi dan perbandingan metode yang pernah digunakan pada kasus serupa.

### Bab III : Metode Penelitian

Bab ini menjelaskan metode yang digunakan dalam penelitian, mulai dari teknik pengumpulan data komentar dari Instagram, proses persiapan data (preprocessing), pemilihan algoritma Transformer IndoBERT untuk klasifikasi sentimen, hingga tahapan evaluasi model. Penelitian ini mengikuti alur CRISP-DM agar proses berjalan sistematis dan terstruktur.

### Bab IV : Hasil dan Pembahasan

Bab ini menyajikan hasil analisis sentimen yang diperoleh dari komentar Instagram terhadap Fakultas Teknik Universitas Dharma Persada. Hasil prediksi sentimen ditampilkan dalam bentuk tabel dan grafik, lalu dianalisis untuk

mengetahui dominasi sentimen positif, netral, atau negatif. Selanjutnya dibahas hasil serta perbandingan dengan penelitian sebelumnya.

#### Bab V : Kesimpulan dan Saran

Bab ini membuat kesimpulan dari hasil penelitian mengenai pola sentimen yang muncul terhadap Fakultas Teknik Universitas Dharma Persada di media sosial Instagram. Selain itu, diberikan saran untuk pihak kampus sebagai bahan pertimbangan perbaikan layanan serta saran bagi penelitian selanjutnya agar hasil yang diperoleh lebih maksimal dan luas cakupannya.

