

BAB I PENDAHULUAN

1.1 Latar Belakang

Telkom University merupakan salah satu perguruan tinggi swasta terbaik di Indonesia yang berfokus pada bidang teknologi, bisnis, dan komunikasi. Dalam rangka memperluas akses pendidikan berkualitas, Telkom University membentuk sejumlah kampus cabang di berbagai daerah di luar kampus utama. Meskipun berada di bawah satu institusi, kampus cabang memiliki karakteristik, kondisi geografis, dan tantangan yang berbeda dari kampus utama.

Dalam konteks pengelolaan reputasi institusi, khususnya kampus cabang, media sosial kini menjadi salah satu saluran penting untuk menangkap opini masyarakat. Internet dan media sosial telah menjadi bagian tak terpisahkan dari kehidupan sehari-hari, menyediakan platform bagi individu untuk berbagi opini tentang berbagai topik, termasuk dunia pendidikan. Masyarakat, khususnya mahasiswa dan calon mahasiswa, sering memanfaatkan media sosial untuk menyampaikan pengalaman, harapan, serta kritik terhadap institusi pendidikan [1]. Ulasan, komentar, dan diskusi yang berkembang di berbagai platform seperti Instagram, Twitter/X, dan TikTok dapat memberikan gambaran mengenai persepsi publik terhadap suatu institusi pendidikan, termasuk kampus cabang Universitas Telkom maupun kampus utamanya.

Dalam bidang pendidikan, analisis sentimen telah terbukti menjadi metode yang optimal untuk memahami persepsi publik serta mendukung pengambilan keputusan berbasis data [2]. Pemanfaatan teknologi *Machine Learning* (ML) dan *Natural Language Processing* (NLP) dalam analisis sentimen memungkinkan institusi pendidikan mengolah data dalam jumlah besar secara optimal, sehingga dapat memberikan wawasan mendalam terkait tingkat kepuasan mahasiswa, efektivitas kebijakan akademik, serta permasalahan yang perlu segera ditindaklanjuti [3].

Meskipun analisis sentimen dalam konteks akademik telah banyak diteliti, sebagian besar penelitian masih berfokus pada institusi pendidikan secara umum atau kampus utama, tanpa mempertimbangkan dinamika di kampus cabang. Padahal, kampus cabang memiliki tantangan tersendiri, seperti persepsi masyarakat terhadap kualitas pendidikan, ketersediaan fasilitas, layanan akademik, serta pengalaman mahasiswa secara keseluruhan. Selain itu, faktor geografis dan demografis dari lokasi kampus cabang dapat memengaruhi cara mahasiswa dan masyarakat menilai institusi tersebut. Persepsi negatif di kampus cabang bahkan berpotensi memengaruhi citra Universitas Telkom secara keseluruhan, sehingga menjadi penting untuk memahami sentimen publik secara spesifik terhadap kampus cabang.

Oleh karena itu, penelitian ini bertujuan untuk mengatasi kesenjangan tersebut dengan menganalisis sentimen publik terhadap kampus cabang Universitas Telkom menggunakan pendekatan NLP dan ML. Dua algoritma yang digunakan dalam penelitian ini adalah Support Vector Machine (SVM) dan Random Forest (RF), yang dipilih berdasarkan reputasinya sebagai metode klasifikasi yang banyak digunakan dalam studi analisis sentimen. Kedua algoritma ini telah terbukti memiliki performa yang baik dalam mengolah data teks dan mampu menghasilkan klasifikasi yang cukup akurat pada berbagai konteks penelitian sebelumnya [4], secara teknis SVM dipilih karena mampu bekerja optimal pada data berdimensi tinggi dan bersifat *sparse* seperti hasil ekstraksi fitur teks dengan TF-IDF. Sementara RF memberikan keunggulan relatif tahan terhadap *overfitting*, sehingga menjadikannya cocok untuk klasifikasi data teks.

Data dikumpulkan dari berbagai platform media sosial untuk memberikan gambaran yang lebih komprehensif mengenai opini masyarakat. Hasil dari penelitian ini diharapkan dapat menjadi dasar bagi Universitas Telkom dalam meningkatkan kualitas akademik, pelayanan, serta strategi komunikasi yang lebih optimal dengan mahasiswa dan masyarakat luas.

Oleh karena itu, penelitian ini tidak hanya menawarkan kontribusi ilmiah dalam ranah analisis sentimen, tetapi juga kontribusi praktis bagi pengembangan mutu pendidikan di kampus cabang Telkom University.

1.2 Rumusan Masalah dan Solusi

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah dan solusi dari proyek akhir ini adalah :

1. Bagaimana membantu manajemen kampus cabang Universitas Telkom dalam mengklasifikasikan sentimen publik terhadap institusi?

Solusi : Menerapkan analisis sentimen berbasis ML dengan algoritma SVM dan RF untuk mengklasifikasikan opini publik dari berbagai platform media sosial ke dalam kategori positif, negatif, atau netral. Hasil klasifikasi ini diharapkan dapat mendukung penyusunan kebijakan kampus yang lebih adaptif terhadap persepsi masyarakat

2. Bagaimana menyajikan hasil klasifikasi sentimen agar dapat memberikan wawasan yang komprehensif bagi manajemen kampus cabang Universitas Telkom?

Solusi : Menyajikan hasil analisis dalam bentuk visualisasi data, seperti *wordcloud* dan tabel kutipan opini yang sudah diberi label sentimen, guna memberikan gambaran yang lebih komprehensif mengenai pola opini publik,

serta membantu perumusan strategi komunikasi dan peningkatan layanan institusi

1.3 Tujuan

Berdasarkan rumusan masalah yang telah dijelaskan, proyek akhir ini bertujuan untuk:

1. Melakukan klasifikasi sentimen publik terhadap kampus cabang Universitas Telkom menggunakan algoritma SVM dan RF, serta mengevaluasi kinerja model berdasarkan metrik evaluasi yang relevan.
2. Membangun *dashboard* interaktif untuk menyajikan hasil klasifikasi sentimen publik terhadap kampus cabang Universitas Telkom sebagai alat bantu analisis opini bagi manajemen universitas.

1.4 Batasan Masalah

Dalam rangka memperjelas ruang lingkup penelitian dan menghindari perluasan topik yang tidak relevan, batasan masalah dalam proyek ini ditetapkan sebagai berikut:

1. Data dikumpulkan dari media sosial publik (Instagram, TikTok, Twitter/X, dan Google Maps) yang secara eksplisit menyebutkan atau membahas kampus cabang Universitas Telkom. Data dikumpulkan melalui teknik web *scraping* baik secara manual maupun otomatis dengan periode pengambilan data antara November 2022 hingga November 2024. Detail dataset yang digunakan dijelaskan dalam Bab III (Analisis Pekerjaan).
2. Opini publik yang dianalisis dibatasi hanya pada opini yang ditulis dalam bahasa Indonesia. Opini yang menggunakan bahasa asing atau campuran tetap dipertimbangkan jika didominasi oleh bahasa Indonesia.
3. Analisis sentimen dibatasi hanya pada kampus cabang Universitas Telkom yang berlokasi di Jakarta, Surabaya, dan Purwokerto. Opini mengenai lokasi lain tidak dianalisis.
4. Klasifikasi sentimen dibatasi pada tiga kategori/Label, yaitu: positif, negatif, dan netral. Sentimen dengan opini yang mengandung lebih dari satu kecenderungan sentimen atau variasi kekuatan emosi akan di pertimbangkan atau tidak dibahas lebih lanjut.
5. Proses ekstraksi fitur untuk data teks dibatasi hanya menggunakan metode Term *Frequency-Inverse Document Frequency* (TF-IDF).

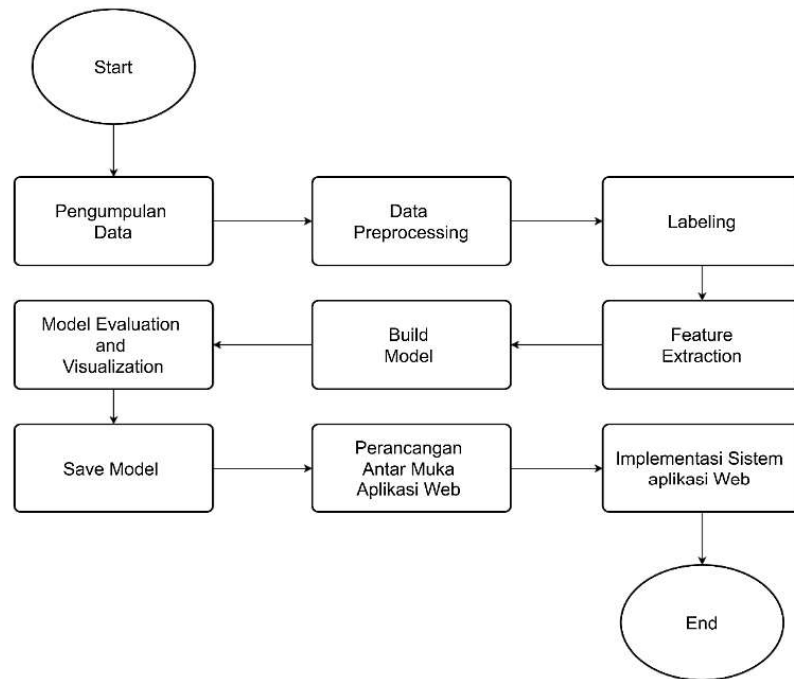
6. Analisis sentimen dilakukan menggunakan dua model ML, yaitu SVM dan RF.
7. Sistem yang dikembangkan dibatasi dalam bentuk aplikasi berbasis web
8. Visualisasi hasil analisis dibatasi pada tampilan tabel opini beserta label sentimennya, serta *wordcloud* untuk masing-masing kategori sentimen

1.5 Metodologi Pengerjaan

Dalam penelitian ini, menggunakan pendekatan metodologi *prototyping* sebagai strategi utama dalam proses pengembangan sistem. Metode ini dipilih karena sifatnya yang fleksibel dan adaptif terhadap perubahan, serta memungkinkan pengembangan dilakukan secara bertahap melalui serangkaian iterasi [5].

Pengembangan dimulai dengan pembuatan prototipe awal yang merepresentasikan fungsi dasar aplikasi, seperti proses analisis sentimen, pembuatan *dashboard*, dan visualisasi data. Prototipe awal tersebut kemudian dievaluasi bersama dengan dosen pembimbing. Berdasarkan hasil evaluasi tersebut, dilakukan revisi dan penyempurnaan terhadap sistem, baik dari sisi fungsionalitas maupun antarmuka pengguna. Siklus ini diulang hingga sistem mencapai stabilitas dan memenuhi tujuan penelitian.

Pendekatan *prototyping* memungkinkan integrasi umpan balik secara berkelanjutan untuk menyempurnakan fitur, meningkatkan performa model ML, serta menyesuaikan struktur data yang digunakan. Dengan demikian, metode ini mendukung pengembangan aplikasi yang responsif terhadap dinamika kebutuhan pengguna dan menghasilkan produk akhir yang representatif serta optimal. Alur pengerjaan dengan pendekatan *prototyping* digambarkan pada Gambar 1.1 berikut:



Gambar 1.1 Alur Pekerjaan

Alur pengerjaan dalam penelitian ini disusun secara sistematis berdasarkan pendekatan metodologi *prototyping*, yang mencakup serangkaian tahapan teknis sebagai berikut:

1. Pengumpulan Data :

Data dikumpulkan dari media sosial publik seperti Instagram, TikTok, Twitter/X, dan Google Maps dengan rentang waktu antara November 2022 hingga November 2024. Teknik pengambilan data dilakukan secara manual maupun otomatis menggunakan *tools* seperti Apify dan *Python scraping*.

2. Data Preprocessing :

Tahapan ini mencakup pembersihan data dari elemen-elemen yang tidak relevan (seperti emoji, URL, dan tanda baca), konversi huruf menjadi huruf kecil, normalisasi kata, serta *stopwords*. Proses ini bertujuan untuk memastikan kualitas data sebelum dianalisis lebih lanjut.

3. Labeling :

Data yang telah diproses kemudian diberi label sentimen (positif, negatif, atau netral) sebagai tahap awal dalam proses *supervised learning*. Proses pelabelan dilakukan secara manual oleh tiga *anotator* berdasarkan konteks komentar yang dianalisis.

4. Feature Extraction :

Komentar yang telah dilabeli dikonversi menjadi representasi numerik menggunakan teknik seperti TF-IDF (Term Frequency–Inverse Document Frequency) agar dapat diproses oleh algoritma ML.

5. Build model :

Model ML dibangun menggunakan algoritma SVM, dan RF, dengan tujuan melakukan klasifikasi sentimen terhadap data yang telah diekstrak fiturnya.

6. Model Evaluation and Visualization :

Model dievaluasi menggunakan metrik evaluasi seperti akurasi, *precision*, *recall*, dan F1-score. Selain itu, hasil klasifikasi divisualisasikan dalam bentuk *confusion matrix* untuk memberikan pemahaman visual terhadap performa model.

7. Save model :

Model yang telah mencapai kinerja terbaik disimpan dan dipersiapkan untuk integrasi ke dalam sistem aplikasi web.

8. Perancangan Antar Muka Aplikasi Web :

Perancangan antarmuka pengguna (UI) dilakukan untuk mendesain tampilan *dashboard* analisis sentimen yang informatif dan interaktif, dengan memperhatikan prinsip *usability* serta kemudahan navigasi.

9. Implementasi Sistem Aplikasi Web :

Sistem aplikasi web diimplementasikan dengan mengintegrasikan model ML ke dalam *backend* aplikasi. Fitur utama yang dikembangkan mencakup unggah data, analisis sentimen, visualisasi hasil, serta pengelolaan riwayat aktivitas pengguna

1.6 Penjadwalan Kerja

Penjadwalan kerja merupakan aspek penting untuk memastikan setiap tahapan proyek berjalan sesuai rencana. Jadwal ini mencakup seluruh aktivitas utama, mulai dari diskusi, pengumpulan dan *preprocessing* data, pelabelan, visualisasi, ekstraksi fitur, pemodelan, evaluasi, penulisan makalah penelitian, revisi proyek, pembuatan dan pengujian *dashboard*, hingga penilaian akhir.

Penyusunan jadwal dilakukan dalam satuan minggu, mencakup periode Agustus 2024 hingga Mei 2025. Setiap minggu dirancang untuk merepresentasikan tahapan pekerjaan yang spesifik, dengan target pencapaian yang terstruktur dan terukur. Rincian lengkap penjadwalan kerja ditampilkan dalam Tabel 1.1 dan Tabel 1.2.

Tabel 1.1 Pelaksanaan Kerja Periode Ganjil 2024/2025

No	Kegiatan	2024																			
		Agustus				September				Oktober				November				Desember			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Diskusi Awal Magang																				
2	Pengumpulan Dataset																				
3	Data Preprocessing																				
4	Labeling																				
5	Visualisasi																				
6	Ekstraksi Fitur																				
7	Modeling																				
8	Evaluasi																				
9	Menulis Paper Penelitian																				

Tabel 1.2 Pelaksanaan Kerja Periode Genap 2024/2025

No	Kegiatan	2025																			
		Januari				Februari				Maret				April				Mei			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	Menulis Paper Penelitian																				
2	Revisi Project																				
3	Pembuatan Dashboard																				
4	Pengujian Dashboard																				
5	Penilaian																				