

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Pada era teknologi saat ini, media *online* dan media sosial telah menjadi platform utama untuk berbagi informasi, pendapat, dan pandangan mengenai berbagai topik, termasuk lembaga pendidikan seperti universitas. Universitas adalah lembaga pendidikan penting dalam masyarakat yang memiliki dampak besar terhadap perkembangan sosial, ekonomi, dan budaya. Opini publik terhadap universitas sangat relevan dalam upaya meningkatkan mutu pendidikan dan citra universitas tersebut.

Salah satu universitas yang aktif dibicarakan saat ini adalah Universitas Telkom (Tel-U). Tel-U merupakan perguruan tinggi swasta yang telah banyak memainkan peran penting dalam pendidikan, inovasi, dan perkembangan teknologi di Indonesia. Menurut data yang dipublikasikan oleh *Webometrics* yang mengelola pemeringkatan 32.000 universitas diseluruh dunia mengungkapkan, pada bulan juli 2023 Tel-U berada di posisi 1.268 universitas terbaik didunia [1]. Hasil pemeringkatan tersebut menjadikan Tel-U sebagai perguruan tinggi swasta terbaik nomor satu di Indonesia. Hal ini dibuktikan dengan berbagai prestasi yang diraih oleh Tel-U, baik di tingkat nasional maupun internasional.

Tel-U memiliki prestasi yang bagus, namun citra dan kepuasan mahasiswa serta masyarakat Indonesia terhadap Tel-U belum tentu sebanding dengan reputasi yang dimilikinya. Beberapa mungkin menyuarakan dukungan dan kepuasan terhadap Universitas Telkom, sementara yang lain mungkin memiliki kritik dan ketidakpuasan terhadap Universitas Telkom. Bukan hanya itu saja banyak dari individu lebih memilih perguruan tinggi negeri dibandingkan perguruan tinggi swasta seperti Tel-U. Adanya variasi dan perbedaan opini ini, diperlukan sebuah analisis terhadap opini publik tersebut. Analisis tersebut menjadi sangat penting bagi Tel-U dalam memperbaiki sistem manajemen dan sistem brandingnya saat ini.

Teknologi yang dapat digunakan untuk menganalisis opini publik adalah *Natural Language Processing* (NLP). NLP adalah sebuah desain dan implementasi sistem komputer untuk menganalisis dan memahami bahasa manusia untuk mengotomatisasi sebuah tugas [2]. Salah satu contoh pengolahan NLP adalah analisis sentimen. Analisis sentimen merupakan salah satu teknik pengolahan NLP yang berguna untuk menganalisis dan mengekstrak informasi terkait opini dan emosi seseorang [3]. Dalam dunia pendidikan, analisis sentimen dapat digunakan untuk mengevaluasi suatu universitas dengan mengkategorikan opini publik mengenai universitas tersebut ke dalam sentimen positif, netral, atau negatif.

Melakukan proses analisis sentimen opini publik terhadap Tel-U, diawali dengan pembangunan model *machine learning* dengan menggunakan algoritma *Support Vector Machine* (SVM). SVM adalah algoritma klasifikasi dan regresi yang menggunakan subset titik yang disebut support vector untuk memisahkan dua kelas besar [4]. Prinsip kerja SVM adalah mencari *hyperplane* (bidang pemisah) yang dapat memisahkan dua kelas data dengan *margin* maksimum [4].

Proyek akhir ini, data opini publik terhadap Tel-U diambil dari hasil *scraping* media *online* dan media sosial. Media *online* yang akan menjadi sumber data analisis ini adalah detik.com dan kompas.com, sedangkan media sosial yang akan menjadi sumber data analisis ini adalah twitter dan instagram. Setelah model *machine learning* berhasil dibangun, model tersebut dapat digunakan untuk mengklasifikasikan sentimen publik terhadap Tel-U yang belum pernah diberi label sentimen sebelumnya.

Menerapkan analisis sentiment pada proyek akhir ini diharapkan akan memberikan pandangan yang lebih mendalam mengenai perasaan dan pandangan publik terhadap Tel-U yang selanjutnya dapat digunakan sebagai panduan strategis yang berharga bagi Tel-U dalam menjaga dan meningkatkan citranya di dunia teknologi yang terus berkembang.

1.2 Rumusan Masalah

Berdasarkan dari latar belakang, rumusan masalah untuk proyek akhir ini adalah:

1. Bagaimana cara membantu manajemen Universitas Telkom (Tel-U) untuk mengklasifikasikan sentimen publik?
2. Bagaimana cara penyajian hasil dari klasifikasi sentimen publik terhadap Universitas Telkom (Tel-U)?

1.3 Tujuan

Berdasarkan rumusan masalah yang telah disebutkan, tujuan yang hendak dicapai dari proyek akhir ini adalah:

1. Menggunakan algoritma *Support Vector Machine (SVM)* untuk merancang sebuah model *machine learning* yang dapat mengklasifikasikan setiap sentimen publik terhadap Universitas Telkom (Tel-U).
2. Membangun dashboard sederhana untuk menyajikan hasil klasifikasi sentimen publik terhadap Universitas Telkom (Tel-U).

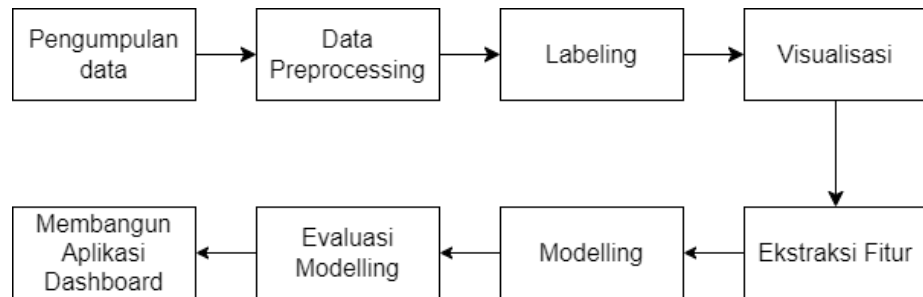
1.4 Batasan Masalah

Beberapa hal yang menjadi batasan masalah dalam proyek akhir ini diantaranya adalah:

1. Data yang digunakan bersumber dari media *online* dan media sosial.
2. Data yang digunakan diambil dari opini publik yang berbahasa Indonesia.
3. Pengambilan data menggunakan metode *scraping*.
4. Hasil klasifikasi berupa sentimen positif, netral dan negatif.
5. Model machine learning yang digunakan adalah algoritma *Support Vector Machine (SVM)*.
6. Hasil analisis sentimen digambarkan dalam bentuk *Wordcloud*.
7. Ekstraksi fitur yang digunakan adalah TF-IDF.

1.5 Metode Pengerjaan

Metode pengerjaan pada dokumen proyek ini terdiri dari beberapa tahap yang dapat dilihat dari Gambar 1. 1 :



Gambar 1. 1 Tahapan Analisis Sentimen

1. **Pengumpulan Data**

Tahap ini merupakan proses pengambilan data. Sumber data dari proyek akhir ini diambil dari media *online* dan media sosial, seperti detik, Kompas, twitter dan instagram.

2. **Data *Preprocessing***

Pada tahap ini data yang telah didapatkan akan dilakukan tahap pembersihan data. Pembersihan data meliputi mengubah format teks menjadi huruf kecil, menghapus *link* tautan, menghapus tanda baca, normalisasi kata, menghapus elemen yang tidak dibutuhkan, memisahkan kalimat menjadi kata perkata (*tokenization*), menghapus kata-kata yang tidak memiliki arti yang signifikan (*stopwords*), mengembalikan kata ke bentuk dasarnya (*stemming*), dan lain sebagainya.

3. **Labeling**

Tahap ini dilakukan proses melabeli data terhadap opini publik apakah opini tergolong sentimen positif, netral, atau negatif.

4. **Visualisasi**

Sentimen publik akan divisualisasikan dalam bentuk *wordcloud*, gunanya untuk menggambarkan hasil polarisasi sentimen data.

5. **Ekstraksi Fitur**

Sebelum melakukan *modeling*, perlu mengubah teks menjadi numerik yang dapat digunakan oleh model. Analisis sentimen ini akan menggunakan metode ekstraksi fitur TF-IDF (*Term Frequency-inverse Document Frequency*).

6. **Modeling**

Pada tahap ini dilakukan pembangunan model *machine learning* menggunakan algoritma SVM. Model ini dibangun dengan melakukan pembagian data yang telah diberi label sentimen pada tahap *labeling*. Data akan dibagi menjadi data *training* 80% dan data *testing* 20%. Data *training* digunakan untuk melatih model dan membuat prediksi, sedangkan data *testing* digunakan untuk menguji performa model yang telah dibangun.

7. **Evaluasi Modeling**

Tahap selanjutnya yaitu evaluasi, di mana dilakukan pengecekan performa model dengan menggunakan *matrix* seperti *accuracy*, *precision*, *recall* dan *F1-score*. Evaluasi ini akan membantu menentukan sejauh mana metode tersebut efektif dalam mengklasifikasikan sentimen. Setelah melakukan evaluasi, *modeling* akan di *save* dan akan digunakan untuk membangun aplikasi dashboard.

8. **Membangun Aplikasi Dashboard**

Tahap terakhir yaitu membuat aplikasi dashboard, dimana model yang telah disimpan nanti akan digunakan didalam aplikasi. Model ini berguna untuk memberi pemahaman ke sistem tentang pola proses analisis sentimen.

1.6 **Jadwal Pengerjaan**

Proyek akhir ini mulai dikerjakan pada bulan oktober 2023 dan berakhir pada bulan juni 2024. Berikut ini jadwal pengerjaan proyek akhir ini yang dapat dilihat di Tabel 1.1 :

Tabel 1. 1 Jadwal Pengerjaan

Kegiatan	2023			2024					
	Okt	Nov	Des	Jan	Feb	Mar	Apr	Mei	Jun
Pengumpulan data	■								
Data Preprocessing		■	■						
Labeling			■	■					
Visualisasi				■					
Ekstraksi Fitur				■	■				
Modeling					■	■			
Evaluasi					■	■			
Membangun dashboard							■	■	
Menulis Paper							■	■	■