

BAB I

PENDAHULUAN

A. Latar Belakang

Pada tahun 2020, Indonesia menempati posisi kedua sebagai penyumbang sampah plastik terbesar di dunia setelah Republik Rakyat Tiongkok, hal ini memunculkan kekhawatiran terkait banyaknya sampah plastik di Indonesia [1]. Berdasarkan data dari Sistem Informasi Pengelolaan Sampah Nasional (SIPSN) pada tahun 2024, tercatat bahwa 41,81% sampah Indonesia belum terkelola, dengan total timbunan lebih dari 18 juta ton per tahun. Meskipun telah terjadi pengurangan sampah sebesar 12,23% di tingkat kabupaten/kota, jumlah sampah yang berhasil dikelola baru mencapai 45,96%. Persebaran fasilitas pengelolaan sampah di Indonesia dilaporkan telah merata sesuai kebutuhan pengelolaan sampah dan jumlah penduduk [2]. Namun, dengan 41,81% sampah yang belum terolah, muncul pertanyaan mengenai tingkat kesadaran masyarakat terhadap pengelolaan dan daur ulang sampah, khususnya plastik.

Media sosial telah menjadi bagian tak terpisahkan dari kehidupan masyarakat Indonesia, dengan pengguna yang terus meningkat dari berbagai kalangan usia dan latar belakang. Popularitas media sosial seperti Facebook, Instagram, Twitter, dan TikTok memberikan peluang besar untuk menyebarkan informasi dan memengaruhi perilaku masyarakat, termasuk dalam mengadopsi praktik berkelanjutan seperti daur ulang. Melalui media sosial, pengguna dapat terinspirasi oleh konten untuk membangun kesadaran dan motivasi untuk ikut berkontribusi menjaga lingkungan [3]. Media sosial tidak hanya berfungsi sebagai alat komunikasi, tetapi juga sebagai media edukasi yang efektif dalam mempromosikan perilaku ramah lingkungan. Dengan tingginya interaksi pengguna, analisis sentimen terhadap opini publik mengenai pengelolaan sampah di media sosial dapat memberikan wawasan penting untuk

meningkatkan kesadaran dan partisipasi aktif dalam kegiatan daur ulang, khususnya daur ulang sampah plastik.

Analisis sentimen terhadap opini publik di media sosial menjadi salah satu pendekatan yang efektif untuk menggali wawasan mendalam mengenai pandangan masyarakat di sosial media terhadap daur ulang plastik. Analisa sentimen adalah proses mengumpulkan dan menganalisis pendapat, pemikiran, dan kesan orang-orang terkait berbagai topik, produk, subyek, dan layanan. Pendapat individu maupun kelompok dapat bermanfaat bagi perusahaan, pemerintah, dan individu lainnya untuk mengumpulkan informasi dan membuat keputusan berdasarkan opini [4].

Sesuai dengan penguraian dari latar belakang, ditemukan bahwa ada sebuah kepentingan untuk mengembangkan sektor daur ulang sampah plastik di Indonesia. Pemahaman umum mengenai daur ulang sampah plastik masih diberikan pandangan samping di Indonesia, ini diakibatkan kurangnya insentif, kurangnya akses terhadap fasilitas pengelolaan sampah, dan persepsi umum yang menganggap bahwa pengelolaan sampah plastik masih kurang optimal. Penggunaan plastik yang berlebih juga menunjukkan pemahaman yang rendah dari masyarakat untuk mengelola sampah plastik lebih lanjut, yang dikombinasikan dengan kurangnya manajemen plastik di Indonesia, menyebabkan masalah sampah plastik di Indonesia [5]. Hasil dari analisis sentimen yang dilakukan pada penelitian diharapkan dapat digunakan untuk menjadi sebuah solusi yang dapat diskalakan untuk mendapatkan opini umum, miskonsepsi, dan pendorong opini yang ada pada masyarakat mengenai daur ulang sampah plastik.

Menurut studi literatur yang dilakukan, masih belum ada *dataset* yang dibuat atau dapat digunakan secara spesifik untuk melakukan analisis sentimen mengenai daur ulang sampah plastik. *Dataset* yang dimaksud adalah *dataset* dari media sosial yang memiliki sentimen secara khusus mengenai daur ulang sampah-sampah plastik beserta jenis-jenisnya dan proses daur ulang yang dapat dilakukan. Seperti daur ulang untuk sampah botol plastik berjenis PET, yang menjadi salah satu masalah untuk perusahaan daur ulang sampah plastik

PET di Indonesia. Apabila pengembangan model analisis aspek berbasis sentimen dilakukan tanpa menggunakan *dataset* yang sesuai, hasil yang didapatkan tidak dapat mencerminkan sentimen sesungguhnya yang dimiliki oleh publik.

B. Rumusan Masalah

Berdasarkan latar belakang yang sudah diuraikan, disusun beberapa rangkaian rumusan masalah sebagai berikut:

1. Bagaimana cara membuat *dataset* sentimen berbasis aspek untuk produk daur ulang plastik dari media sosial Twitter dan Facebook?
2. Bagaimana cara membangun model pemrosesan bahasa alami (*Natural Language Processing/NLP*) yang efektif untuk mengklasifikasikan sentimen berbasis aspek dengan pendekatan *multi-label* secara akurat berdasarkan data yang telah diproses?

C. Batasan Masalah

Untuk melakukan penelitian, ada beberapa batasan yang harus ditetapkan untuk rumusan masalah yang telah ditetapkan agar penelitian dapat terarah dan tertuju dengan analisa sentimen mengenai daur ulang sampah plastik.

1. Pengambilan data adalah menggunakan metode *web scraping* dan menggunakan API untuk sosial X (Twitter) dan Facebook.
2. Keseluruhan data yang diambil berupa tulisan/judul/komen dari unggahan yang dibuat secara publik. Data pribadi maupun unggahan privat tidak diambil dalam penelitian.
3. Data yang digunakan adalah data yang mayoritas (>90%) berbahasa Indonesia.
4. Aspek yang digunakan untuk pembangunan *dataset* adalah *recyclability of plastic in general, recyclability of PET plastic, recycling process, dan future of recycling*.

5. Model-model yang dibuat akan dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-Score*.

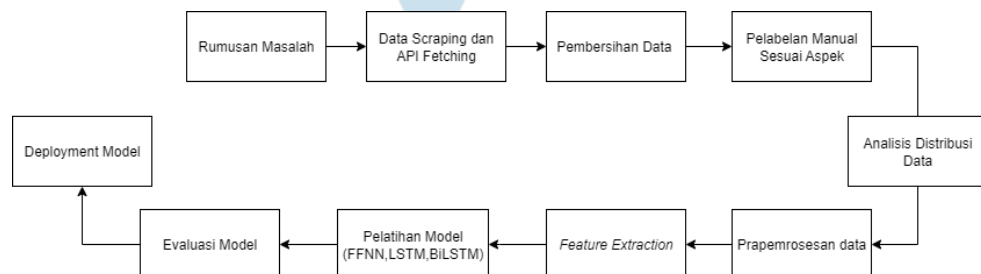
D. Tujuan Penelitian

Sesuai dengan rumusan masalah yang telah dibuat sebelumnya, ada beberapa tujuan yang harus dicapai untuk keberhasilan dari penelitian ini. Tujuan dari penelitian pembuatan analisa sentimen untuk daur ulang plastik berdasarkan data dari sosial media adalah sebagai berikut:

1. Membuat *dataset* yang dapat digunakan secara spesifik mengenai daur ulang sampah plastik dengan menggunakan data dari media sosial Twitter dan Facebook.
2. Membangun model pemrosesan bahasa alami (*Natural Language Processing/NLP*) untuk klasifikasi sentimen berdasarkan data yang telah diproses.

E. Metode Penelitian

Penelitian akan dilakukan dengan proses yang digunakan pada umumnya untuk pembuatan model pembelajaran mesin yaitu dengan proses seperti pada Gambar 1.1. Proses yang dilakukan meliputi pengumpulan data dalam bentuk *web scraping* dan penggunaan API, persiapan data untuk digunakan pada pelatihan model, pelatihan dan evaluasi model, dan berakhir pada *deployment* untuk model terbaik. Persiapan data pada penelitian mencakup pengumpulan, pembersihan, dan pelabelan data untuk pembangunan *dataset*.



Gambar 1.1. Proses Penelitian dan Pengembangan

Pengumpulan data dilakukan dengan menggunakan teknik *web scraping* dan menggunakan API untuk mendapatkan dari media sosial Twitter

dan Facebook. Pengumpulan data diiringi dengan pembersihan data untuk menghapus data duplikat dan menghilangkan bagian-bagian dari data yang tidak digunakan pada pengembangan model, seperti kolom-kolom yang tidak relevan dan *link* pada teks.

Pelabelan data dilakukan secara manual dengan bantuan dari sebuah aplikasi, yaitu *label-studio* yang berfungsi sebagai sebuah antarmuka untuk menampilkan dan melabel data dengan fungsi *import* dan *export* untuk memudahkan pemindahan data. Analisis data dilakukan setelah data secara keseluruhan sudah dilabel, dengan analisis data, dapat ditemukan apabila data yang didapatkan memiliki ketimpangan pada aspek-aspek dan apabila dibutuhkan pengumpulan data lebih lanjut atau menggunakan teknik-teknik untuk menyeimbangkan data.

Data yang sudah dianalisis dan dikumpulkan dengan fitur *export* akan diintegrasikan untuk digunakan pada pelatihan model. *Preprocessing* data lalu dilakukan untuk mengubah data menjadi bentuk yang dapat digunakan untuk proses pelatihan dan evaluasi model. Data akan dibagi menjadi data pelatihan dan data evaluasi, lalu diubah dari teks menjadi *token* agar dapat dimengerti oleh model.

Pelatihan model dilakukan untuk tiga jenis model, yaitu model *feed forward neural network*, model LSTM, dan model BiLSTM. Pelatihan model akan dilakukan untuk mencari model terbaik, dengan model *feed forward neural network* menjadi sebuah model *baseline* untuk dibandingkan pada LSTM dan BiLSTM. Dari ketiga model yang dilatih dengan data, akan dilakukan evaluasi dengan menggunakan metrik *accuracy*, *precision*, *recall*, dan F1-Score serta menggunakan *confusion matrix* untuk visualisasi kemampuan model pada *dataset* evaluasi.

F. Sistematika Penulisan

Sistematika penulisan tugas akhir yang telah dibuat oleh peneliti akan terdiri dari beberapa bab yaitu sebagai berikut:

BAB I Pendahuluan

Bab pertama pada tugas akhir berisikan mengenai latar belakang masalah, rumusan masalah, batasan penelitian, tujuan penelitian, metode penelitian yang digunakan pada penelitian, dan sistematika penulisan.

BAB II Tinjauan Pustaka.

Bab kedua pada tugas akhir berisi penelitian-penelitian terdahulu yang memiliki kemiripan topik dengan penelitian yang dilakukan oleh peneliti. Sumber dari penelitian-penelitian yang digunakan adalah skripsi, jurnal, maupun artikel yang didapatkan dari sumber yang bertanggung jawab dan dapat diverifikasi kebenarannya.

BAB III Landasan Teori

Bab ketiga pada tugas akhir berisi penjelasan mengenai teori-teori dan kutipan yang digunakan pada penelitian untuk membantu agar pembaca dapat memahami istilah-istilah yang digunakan oleh penulis. Penjelasan yang diberikan atas teori-teori maupun kutipan didapatkan dari sumber-sumber, seperti buku atau jurnal, yang bertanggung jawab.

BAB IV Dataset dan Pengembangan Model.

Bab keempat pada tugas akhir menjelaskan mengenai *dataset* yang digunakan untuk melakukan pengembangan dan juga penjelasan mengenai pengembangan model yang dilakukan untuk penelitian.

BAB V Implementasi Model dan Pengujian Sistem

Bab kelima pada tugas akhir ini adalah penjelasan implementasi dari model terbaik pada sistem, serta pengujian dari sistem tersebut.

BAB VI Penutup

Bab terakhir pada tugas akhir ini berisi kesimpulan dan saran dari peneliti terkait dengan penelitian yang dilakukan untuk kedepannya bisa dikembangkan lagi dengan lebih baik.