

BAB I

PENDAHULUAN

1.1. Latar Belakang

Dalam era digital yang mengalami perkembangan yang sangat cepat dan pengguna perangkat elektronik yang terus bertambah, berita dan informasi terus mengalir dengan sangat cepat dan bertumbuh dalam jumlah yang sangat besar. *Platform online* telah menjadi tempat dimana orang-orang menyampaikan opini dan berbagi pengalaman mereka terhadap hal-hal yang terjadi di dalam kehidupan mereka dan hal-hal yang mereka rasakan. *Google Play Store* menjadi salah satu *platform online* dimana para pengguna perangkat elektronik menyampaikan opini dan berbagi pengalaman mereka dalam menggunakan produk atau aplikasi yang tersedia. Mereka memberikan ulasan serta rating terhadap aplikasi yang mereka gunakan, ulasan yang mereka berikan menjadi cerminan sentimen mereka yang dapat berupa sentimen positif, negatif ataupun netral. Tidak hanya sekedar berbagi opini dan pengalaman saja, para pengguna aplikasi juga secara tidak langsung memberikan wawasan mengenai kualitas dari aplikasi yang mereka gunakan.

Dengan meningkatnya transformasi digital, data telah menjadi salah satu sumber daya yang sangat berharga dan menjadi faktor penentu dalam pengambilan keputusan bisnis. Bagi perusahaan yang bergerak di bidang teknologi dan sekaligus menjadi pengembang aplikasi seperti META, yang menjadi pemilik salah satu aplikasi yang paling populer dan paling banyak digunakan di seluruh dunia seperti Instagram, memahami sudut pandang dan sentimen para pengguna merupakan faktor kunci dalam mempertahankan loyalitas pengguna untuk terus menggunakan

aplikasi mereka serta melakukan peningkatan kualitas untuk aplikasi mereka. Sentimen yang positif dapat memberikan dampak yang positif juga seperti peningkatan reputasi aplikasi dan membantu dalam meningkatkan jumlah unduhan. Sementara itu sentimen negatif dapat memberikan dampak yang negatif bagi perusahaan dengan meningkatnya resiko penurunan reputasi dan jumlah pengguna yang menggunakan aplikasi mereka, tetapi hal itu tidak akan terjadi apabila perusahaan META menerima ulasan atau sentimen negatif yang diberikan pengguna sebagai *feedback* yang konstruktif untuk perbaikan lebih lanjut terhadap aplikasi mereka. Namun, dengan begitu banyaknya ulasan yang diterima setiap hari, semakin sulit untuk mengkurasi dan memahami sentimen dari data tersebut secara manual.

Dengan teknologi *deep learning* seperti pemrosesan bahasa alami (*Natural Language Processing*) pengolahan data sentimen secara otomatis kini dapat dilakukan dengan lebih presisi dan efisien. Pembelajaran mendalam atau *Deep learning* berkembang sebagai cabang bagian dari *machine learning* yang menggunakan jaringan saraf tiruan (*artificial neural networks*) yang dimana pada jaringan saraf tiruan tersebut terdapat banyak lapisan (*layers*) yang disebut *multilayered* untuk melakukan ekstraksi fitur dari data yang di *Input* (Bishop and Bishop 2023). Dengan menggunakan *deep learning* komputer dapat lebih memungkinkan mempelajari representasi data kompleks pada data teks, suara ataupun gambar.

Dalam melakukan pekerjaan seperti analisis sentimen, terdapat salah satu model NLP yang memiliki kinerja yang cukup baik, yaitu model *RoBERTa*. Model *RoBERTa* sebagai pengembangan dari model BERT (Bidirectional *encoder* representations from transformers), yang di latih dengan menggunakan *dataset* yang jauh lebih besar sekitar 10 kali lebih besar dari *dataset* yang digunakan untuk pelatihan model BERT dan dilatih dengan waktu pelatihan yang lebih lama (Kurniadi et al. 2024). *RoBERTa* dibangun dengan menggunakan arsitektur dari algoritma *Transformers* yang dirancang khusus melakukan tugas pengenalan dan pemrosesan bahasa alami manusia seperti analisis sentimen. Model *RoBERTa* juga memiliki variasi model untuk melakukan tugas pemrosesan bahasa alami khusus untuk bahasa Indonesia yaitu *Indonesian-RoBERTa*. *Indonesian-RoBERTa* model yang dibangun dari model *RoBERTa* yang dilatih dengan dataset SmSA, dataset ini berisi kalimat-kalimat opini berbahasa Indonesia yang telah diberi label sentimen, yaitu positif, netral atau negatif. Model ini mampu memahami konteks kalimat dalam bahasa Indonesia dengan lebih baik dibandingkan model lain yang umum digunakan, seperti SVM atau Naive Bayes. Untuk mendapatkan sentimen ulasan para pengguna instagram di *Google Play Store* penulis membuat sebuah program *web scraping* dengan bahasa pemrograman python, teknik *web scraping* digunakan untuk melakukan pengumpulan data dari *Google Play Store* secara otomatis. Dengan teknik ini, nantinya data ulasan para pengguna aplikasi Instagram yang terdapat di *Google Play Store* dapat dikumpulkan dan dianalisis secara serentak.

Masalah umum dalam penelitian ini mencakup upaya mengotomatisasi pengumpulan data sentimen pengguna aplikasi instagram yang mencapai ratusan bahkan ribuan data di *Google Play Store*, serta bagaimana cara memahami sentimen dari setiap ulasan yang diberikan oleh pengguna. Selain itu, dengan semakin banyaknya data ulasan pengguna instagram yang sudah beredar di *Google Play Store*, tentunya akan semakin sulit untuk mengkurasi dan memahami sentimen dari seluruh ulasan tersebut secara manual.

Berdasarkan permasalahan yang terdapat pada penelitian ini, dan untuk mengatasi permasalahan yang terjadi, membangun analisis sentimen *dashboard* menjadi solusi yang cukup efektif untuk dapat memahami setiap emosi dan opini yang tersembunyi di balik sentimen para pengguna. Oleh karena itu, penulis melakukan penelitian ini bertujuan untuk mengembangkan dan mengimplementasikan model *Indonesian-RoBERTa* untuk membangun sebuah *dashboard* analisis sentimen yang dapat mengumpulkan dan menganalisis ulasan aplikasi Instagram dari *Google Play Store*. *Dashboard* ini diharapkan dapat memberikan wawasan yang lebih dalam mengenai sentimen pengguna dan membantu dalam pengambilan keputusan yang lebih baik.

1.2. Identifikasi masalah

Berdasarkan latar belakang yang telah di uraikan, terdapat beberapa permasalahan yang dapat diidentifikasi sebagai berikut:

1. Kesulitan dalam mengumpulkan data sentimen dari aplikasi Instagram yang ada di *Google Play Store*.

2. Tantangan dalam mengimplementasikan model Indonesian *RoBERTa* kedalam *dashboard* analisis sentimen.
3. Melakukan perancangan sebuah *dashboard* yang efektif untuk menampilkan hasil analisis sentimen pengguna aplikasi Instagram.

1.3. Batasan masalah

Berdasarkan permasalahan yang sudah di uraikan, terdapat beberapa batasan masalah yang menjadi batasan penulis dalam melakukan penelitian ini:

1. Penelitian ini hanya akan mengumpulkan dan menganalisis data sentimen dari aplikasi Instagram yang terdapat di *platform Google Play Store*.
2. Pada penelitian ini model *Indonesian-RoBERTa* hanya akan digunakan untuk menganalisis teks dalam bahasa Indonesia, bahasa lain tidak akan dianalisis
3. Penelitian ini tidak akan mencakup aspek keamanan dan privasi data, fokus utama penulis adalah pada teknik pengumpulan data dan analisis sentimen.

1.4. Rumusan Masalah

dari latar belakang yang telah di uraikan, terdapat beberapa masalah yang dapat diidentifikasi sebagai berikut:

1. Bagaimana cara mengumpulkan data ulasan para pengguna aplikasi instagram yang terdapat di *Google Play Store*?
2. Bagaimana mengimplementasikan model *Indonesian-RoBERTa* kedalam *dashboard* analisis sentimen yang akan dibangun?
3. Bagaimana merancang dan membangun sebuah *dashboard* yang efektif untuk menampilkan hasil analisis sentimen dari aplikasi instagram?

1.5. Tujuan Penelitian

Penelitian ini memiliki beberapa tujuan utama yang diharapkan dapat dicapai.

Tujuan-tujuan tersebut adalah sebagai berikut:

1. Untuk mengembangkan metode yang efisien untuk mengumpulkan data sentimen dari aplikasi Instagram yang terdapat di *Google Play Store* menggunakan teknik *web scraping*.
2. Untuk mengimplementasikan model *Indonesian-RoBERTa* ke dalam *dashboard* analisis sentimen dengan tujuan untuk mengklasifikasikan sentimen dari data ulasan pengguna dengan akurat.
3. Untuk merancang dan membangun *dashboard* analisis sentimen yang efektif untuk menampilkan hasil analisis sentimen dari aplikasi Instagram di *Google Play Store*.

1.6. Manfaat penelitian

Penelitian ini diharapkan memberikan berbagai manfaat bagi berbagai pihak, baik secara langsung maupun tidak langsung. Berikut adalah beberapa manfaat yang dapat diperoleh dari penelitian ini.

1.6.1. Manfaat Teoritis:

1. Penelitian ini bertujuan mengembangkan pengetahuan tentang metode analisis sentimen dengan model *Indonesian-RoBERTa* dan *web scraping* untuk pengumpulan data secara otomatis.

2. Penelitian ini bertujuan memberikan kontribusi akademis dengan menghasilkan referensi yang bermanfaat bagi penelitian selanjutnya terkait analisis sentimen.
3. Penelitian ini bertujuan mendorong pengembangan metode analisis sentimen yang lebih baik untuk penelitian selanjutnya, dengan menawarkan wawasan baru tentang penggunaan model NLP dalam pengolahan data sentimen.

1.6.2. Manfaat Praktis

1. Penelitian ini bertujuan menyediakan *tools* yang membantu orang-orang untuk memahami sentimen umum mengenai aplikasi Instagram dengan lebih akurat melalui tampilan *dashboard*.
2. Penelitian ini bertujuan meningkatkan efisiensi dalam mengelola dan menganalisis sentimen pengguna, sehingga dapat membuat keputusan yang lebih cepat dan tepat berdasarkan hasil analisis.
3. Penelitian ini bertujuan mendukung pengembangan keterampilan teknis penulis dalam penerapan model *NLP*, serta perancangan dan pembangunan *dashboard* untuk aplikasi analisis sentimen.