

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Dalam era digitalisasi yang semakin maju, *e-commerce* dan media sosial telah mengalami perkembangan signifikan dan memainkan peran penting dalam transformasi bisnis dan interaksi masyarakat. Berdasarkan data Statista Market Insights, pada tahun 2023 pengguna *e-commerce* di Indonesia mencapai angka 196,47 juta, di mana angka tersebut meningkat sebesar 9,78% dari tahun 2022 [1]. Adapun Databoks pada 2023 menghimpun data dari 5 aplikasi *e-commerce* di Indonesia yang paling banyak dikunjungi, di mana Tokopedia menduduki posisi teratas dengan 1,25 miliar total pengunjung. Sementara itu, pengguna media sosial secara keseluruhan di Indonesia mencapai angka 167 juta pada 2023, di mana jumlah tersebut setara dengan 60,4% populasi domestik [2]. Beberapa sosial media yang umum digunakan di Indonesia antara lain Facebook, Instagram, Twitter, YouTube dan TikTok. Diantara kelima media sosial tersebut, TikTok adalah media sosial yang paling pesat mengalami pertumbuhan pengguna. Adapun TikTok secara signifikan mengalami pertumbuhan pengguna per Juli 2024 sebanyak 157,6 juta, di mana apabila dikalkulasikan dari jumlah pengguna pada Oktober 2023 yang berjumlah 106,2 juta, maka kenaikan pengguna TikTok berada di angka 48,4% [3]. Hal tersebut dapat dipengaruhi oleh berbagai faktor, utamanya adalah platform TikTok itu sendiri yang bukan hanya sebagai sosial media tetapi juga memiliki fitur *e-commerce* di dalamnya, peleburan platform ini disebut sebagai *social commerce*.

Berdasarkan data yang sudah disebutkan, dapat diketahui bahwa pertumbuhan atau peningkatan pengguna *e-commerce* dan media sosial berbanding lurus. Hal tersebut mengindikasikan pergeseran signifikan dalam perilaku konsumen yang semakin mengadopsi model bisnis digital. Pertumbuhan yang pesat dalam penggunaan TikTok, menunjukkan bahwa platform tersebut bukan hanya tempat untuk berbagi konten hiburan, tetapi juga menjadi platform yang menarik untuk aktivitas *e-commerce* melalui konsep *social commerce*.

Pada 12 Desember 2023, TikTok yang beroperasi di bawah naungan ByteDance secara resmi mengakuisisi 75% saham Tokopedia melalui GoTo (Gojek Tokopedia) dengan nilai 1,5 miliar USD atau sekitar 23,5 triliun Rupiah. Dalam

akuisisi ini, TikTok akan memiliki kontrol atas PT Tokopedia dalam kemitraan strategis antara TikTok dan GoTo. Sementara itu PT Tokopedia akan mengelola dan mengoperasikan fitur belanja pada aplikasi TikTok di Indonesia. Pernyataan resmi menyebut investasi TikTok sebagai komitmen jangka panjang untuk mendukung operasional Tokopedia. Namun, ada kekhawatiran terkait dominasi pasar, persaingan tidak sehat, dan keamanan data pengguna [4]. Spekulasi bermunculan terkait akuisisi TikTok pada Tokopedia yang mana hal tersebut dinilai hanya sebagai cara curang TikTok untuk mengantongi izin dagang di Indonesia. Dalam konteks tersebut, TikTok Shop di Indonesia pernah ditutup karena masalah perizinan pada tanggal 4 Oktober 2023 lalu. Menteri Koperasi dan UKM, Tetan Masduki, menjelaskan bahwa izin TikTok hanya sebagai kantor perwakilan perusahaan perdagangan asing, sehingga TikTok hanya dapat memfasilitasi promosi, bukan bertransaksi secara langsung melalui *e-commerce*. Hal ini sesuai dengan revisi Peraturan Menteri Perdagangan Nomor 31 tahun 2023 yang membedakan *social commerce* dengan *social media*. Meskipun terdapat juga potensi keuntungan dari akuisisi ini, seperti dukungan bagi UMKM melalui program TikTok dan Tokopedia, pemasaran, *branding*, dan praktik bisnis berkelanjutan. Tetap tidak dapat dipungkiri adanya potensi monopoli dan pengaruh geopolitik dari China. TikTok sendiri telah menjadi kontroversial karena dinilai sebagai alat propaganda pemerintahan China, banyak negara yang sudah melakukan blokir pada platform ini seperti Amerika Serikat, Kanada, hingga India dengan ancaman kedaulatan sebagai alasan utama [5]. Maka dari itu, akuisisi ini memunculkan pertanyaan tentang kepemilikan data, keamanan data pengguna, dan geoekonomi di Indonesia.

Dalam konteks akuisisi TikTok terhadap Tokopedia melalui GoTo, analisis sentimen menjadi penting untuk dilakukan. Analisis sentimen adalah proses untuk mengevaluasi apakah data memiliki sikap positif, negatif, atau netral [6]. Analisis sentimen akan memungkinkan kita untuk menggali lebih dalam tentang bagaimana masyarakat menangani dan merasakan perubahan ini. Sebagai contoh, apakah mereka mempertimbangkan akuisisi ini sebagai kesempatan atau ancaman? Apakah pendapat mereka di media sosial mencerminkan kekhawatiran tentang dominasi pasar dan persaingan yang tidak sehat? Melalui analisis ini, kita dapat

mengidentifikasi pola sikap dan pandangan yang dapat membantu perusahaan, regulator, dan masyarakat untuk memahami dampak sosial dan ekonomi dari akuisisi yang sudah dilakukan.

Banyak penelitian yang sudah dilakukan terkait analisis sentimen berdasarkan teks emosi. Seperti dalam penelitian yang berjudul “*Attention-Based Modeling For Emotion Detection And Classification In Textual Conversations* [7].” Penelitian ini melibatkan pengembangan model berbasis pembelajaran untuk mengidentifikasi emosi dalam percakapan teks. Dataset mereka mencakup enam jenis emosi berdasarkan Paul Ekman. Setelah data dikumpulkan, data ditokenisasi dan dimasukkan ke dalam *encoder*. Selanjutnya, *encoder* mengirimkan informasi ke unit Bi-LSTM yang dilatih dengan *average stochastic gradient descent* (ASGD). Mekanisme *dropout* digunakan untuk mencegah *overfitting*. Penelitian juga mengimplementasikan mekanisme *self-attention* untuk identifikasi emosi dalam percakapan. Klasifikasi data ke dalam kategori emosi yang sesuai dilakukan menggunakan lapisan *dense* dan aktivasi SoftMax, di mana hasil F1 score mencapai angka 75,82%. Kemudian penelitian dari [8] dengan judul “*Sentiment Classification Using Convolutional Neural Networks*” yang mana dilakukan perancangan sebuah arsitektur CNN untuk klasifikasi sentimen pada data teks pada tiga dataset yang berbeda yaitu dataset IMDB, Yelp, dan Amazon. Dalam penelitian tersebut, dilakukan juga proses *embedding* untuk mengubah kata menjadi vektor numerik. Dari *layer embedding*, dilanjutkan pada pembuatan layer konvolusi, *pooling*, dan *fully connected* pada model CNN. Setelah perbandingan model dilakukan, hasil eksperimen menunjukkan bahwa model CNN yang dirancang dalam penelitian memiliki performa yang lebih baik dibandingkan dengan model *deep learning* lainnya dengan akurasi 90,1% pada dataset IMDB, 87,2% pada dataset Yelp, dan 83,8% pada dataset Amazon. Selanjutnya penelitian dengan judul “*An Efficient CNN-LSTM Model for Sentiment Detection in #BlackLivesMatter* [9]” Di mana dilakukan penggabungan model dari CNN dengan LSTM untuk melakukan analisis sentimen terhadap *hashtag black lives matters* pada dua provinsi di Amerika. Dengan menggabungkan kelebihan CNN dalam pengenalan fitur spasial dan kelebihan LSTM dalam pemahaman konteks, model CNN-LSTM dapat secara efektif menangkap informasi penting dari teks yang diperlukan dalam

analisis. Penelitian tersebut membandingkan model CNN-LSTM yang dirancang dengan beberapa model lain seperti Random Forest dan BERT, yang mana model CNN-LSTM yang diusulkan mencapai tingkat akurasi tertinggi sebesar 94% dalam mendeteksi berbagai sentimen pada kasus terkait.

Berdasarkan permasalahan yang sudah disebutkan dan penelitian terdahulu, *text-based emotion* atau emosi berdasarkan teks dengan mengadopsi metode gabungan *deep learning* LSTM dan CNN digunakan dalam penelitian ini sebagai pendekatan inovatif untuk menggali dan menganalisis dinamika emosional dalam data teks. Dengan melakukan identifikasi terhadap emosi perasaan seperti kegembiraan, kesedihan, kemarahan, kebahagiaan, kebencian, ketakutan, dan lainnya, suatu masalah dapat ditunjukkan dengan jelas. Dengan pemahaman emosi di balik teks, konteks data yang diperoleh akan menjadi lebih kaya daripada hanya mengenali sentimen positif, negatif, atau netral. Adapun dalam penelitian ini, komentar pengguna dari platform media sosial YouTube pada topik terkait akan digunakan sebagai data primer untuk diolah kemudian dilakukan analisis sentimen berdasar teks emosi dengan menggunakan *Plutchik emotion model*.

Analisis sentimen dan emosi memiliki berbagai cara pengaplikasian serta dapat dilakukan dengan menggunakan beragam metodologi. Terdapat tiga jenis teknik analisis sentimen dan emosi: *lexicon based*, *machine learning based*, dan *deep learning based*. Masing-masing metodologi memiliki keunggulan dan kekurangan dalam penerapannya. Dalam penelitian ini, pendekatan *deep learning* akan digunakan. *Deep learning* sendiri merupakan konsep umum yang mencakup varian-varian baru dari sejumlah model pembelajaran yang sudah ada, *deep learning* lebih dikenal dengan istilah *neural networks*, yang kini lebih umum disebut DNN (*Deep Neural Networks*) [10]. *Deep learning* memiliki keunggulan dalam menangani data kompleks dan non-linear seperti teks berbahasa natural. Model *deep learning*, terutama dalam bentuk *neural networks*, dapat mengekstrak representasi fitur yang mendalam dan kompleks dari teks [11], yang memungkinkan mereka untuk mengenali dan memahami nuansa emosi dengan tingkat akurasi yang lebih tinggi. Kemudian *deep learning* juga mampu menangkap fitur sintaksis (seperti urutan kata) dan semantik (makna) dalam teks. Ini membantu meningkatkan akurasi dalam memprediksi sentimen. Dalam konteks analisis

sentimen berbasis teks emosi, dengan bantuan dari natural *language processing* (NLP) untuk memproses data teks dari komentar pada YouTube, digunakan LSTM (*Long-short Term Memory*) dan CNN (*Convolutional Neural Networks*) sebagai pemodelan *deep learning*. Model LSTM sendiri merupakan varian dari *recurrent neural networks* (RNN). LSTM membantu mempertahankan kesalahan yang dapat dipropagasi mundur melalui waktu dan layer yang dimiliki [12]. Hal tersebut memungkinkan RNN untuk dapat melakukan *learning* data secara kontinu selama waktu yang dibutuhkan dengan menjaga error tetap konstan. Kemudian CNN, sering disebut ConvNet, memiliki *deep feed-forward* arsitektur dengan kemampuan generalisasi yang baik. Kepentingan utama dalam menerapkan CNN terletak pada ide penggunaan konsep pembagian bobot, yang mengakibatkan jumlah parameter yang perlu dilatih secara signifikan berkurang, menghasilkan generalisasi yang lebih baik. Dengan jumlah parameter yang lebih sedikit, CNN dapat dilatih dengan lebih mudah dan dapat menghindari *overfitting* data [13].

Berdasarkan uraian di atas, analisis sentimen berbasis teks emosi pada YouTube dengan pendekatan *deep learning* dalam akuisisi TikTok terhadap Tokopedia bertujuan untuk mengetahui sentimen pengguna pada topik terkait dengan melibatkan aspek emosi untuk dianalisis secara komprehensif menggunakan pembelajaran *deep learning* pada prosesnya. Sehingga perusahaan, masyarakat, dan pembuat kebijakan terkait, dapat memahami dampak sosial dan ekonomi akibat akuisisi TikTok terhadap Tokopedia melalui identifikasi pola sikap yang dilakukan. Pemodelan CNN dan LSTM diharapkan dapat memberikan representasi fitur yang mendalam, memungkinkan pemahaman yang lebih akurat terhadap nuansa emosional dalam komentar pengguna.

## **1.2. Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah yang dapat diangkat dalam penelitian ini adalah sebagai berikut:

1. Bagaimana proses penerapan *deep learning* dalam analisis sentimen dan emosi berbasis teks pada YouTube dengan topik akuisisi TikTok terhadap Tokopedia dilakukan?

2. Bagaimana hasil analisis sentimen dan emosi berbasis teks pada YouTube dengan topik akuisisi TikTok terhadap Tokopedia menggunakan pendekatan *deep learning*?

### **1.3. Batasan Masalah**

Bersasarkan perumusan masalah yang sudah disebutkan, berikut adalah batasan masalah dalam penelitian ini:

1. Data yang diambil bersumber dari komentar video YouTube yang membahas akuisisi TikTok terhadap Tokopedia pada rentang waktu 11 Desember 2023 – 20 Februari 2024.
2. Pengambilan data dilakukan secara *scraping* menggunakan YouTube API.
3. Bahasa pemrograman yang digunakan dalam memproses data adalah Python.

### **1.4. Tujuan Penelitian**

Berdasarkan uraian yang telah dipaparkan, berikut adalah tujuan yang ingin dicapai dalam penelitian ini:

1. Mengetahui bagaimana pendekatan *deep learning* dapat diimplementasikan dalam melakukan analisis sentimen dan emosi berbasis teks pada YouTube dengan topik akuisisi TikTok terhadap Tokopedia.
2. Mengetahui bagaimana hasil analisis sentimen dan emosi berbasis teks pada YouTube dengan topik akuisisi TikTok terhadap Tokopedia menggunakan pendekatan *deep learning*.

### **1.5. Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan gambaran yang lebih baik tentang bagaimana masyarakat merespons akuisisi TikTok terhadap Tokopedia, sehingga dapat menjadi pedoman bagi perusahaan bisnis, masyarakat, dan pembuat kebijakan untuk evaluasi dampak yang ditimbulkan.

2. Penggunaan *deep learning*, termasuk LSTM dan CNN, memungkinkan identifikasi pola emosional kompleks dalam data teks, memberikan pemahaman akurat mengenai sikap dan pandangan masyarakat terhadap topik yang dianalisis.
3. Memberikan kontribusi pada pengembangan metodologi analisis sentimen dan emosi berbasis teks dengan pendekatan *deep learning*, yang dapat bermanfaat dalam konteks serupa untuk penelitian lebih lanjut.

### 1.6. Sistematika Penulisan

Sistematika penulisan penelitian membantu dalam mengarahkan penyusunan penelitian agar tidak terdapat penyimpangan sekaligus sebagai acuan untuk mencapai tujuan penulisan sesuai dengan apa yang diharapkan. Langkah-langkah dalam proses penyusunan penelitian ini adalah sebagai berikut:

#### **BAB I**

#### **PENDAHULUAN**

Berisi mengenai gambaran umum dari isi penelitian diantaranya latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat dan sistematika penulisan.

#### **BAB II**

#### **TINJAUAN PUSTAKA**

Berisi mengenai pengertian dasar teori yang berkaitan dengan permasalahan yang dibahas, metode, studi literatur, hingga *tools* pada penelitian.

#### **BAB III**

#### **METODOLOGI PENELITIAN**

Berisi mengenai langkah-langkah yang dilakukan untuk mencapai tujuan penelitian yang mana meliputi identifikasi masalah, metode pengumpulan data, pemrosesan data, dan penulisan laporan.

**BAB IV****PEMBAHASAN**

Berisi mengenai hasil dari setiap tahapan yang terdapat pada metodologi penelitian dan pembahasan terkait proses pengumpulan data, melabeli data, tahapan *preprocessing*, modeling, hingga evaluasi model.

**BAB V****KESIMPULAN**

Berisi mengenai kesimpulan yang dapat diambil dari keseluruhan isi yang berkorelasi dengan rumusan masalah penelitian serta saran untuk pengembangan model atau terkait dengan pemrosesan data untuk penelitian selanjutnya.

**DAFTAR PUSTAKA**

Berisi mengenai literatur yang digunakan sebagai pedoman yang membantu pengerjaan penelitian.