BAB 1

PENDAHULUAN

1.1 Latar Belakang

Seiring dengan kemajuan teknologi, masyarakat telah memanfaatkan media sosial sebagai medium penting untuk berkomunikasi serta menyampaikan opini secara instan. Platform seperti X memungkinkan warganet untuk berbagai pandangan tentang berbagai isu, baik positif maupun negatif, yang sering kali memicu diskusi luas [1]. Fitur seperti batas karakter 280 karakter, *retweet* instan, dan penggunaan tagar membuat isu-isu terkini cepat tersebar dan menarik perhatian publik [2]. Dalam konteks ini, analisis sentimen berfungsi sebagai instrumen penting dalam menelaah pandangan masyarakat terhadap isu-isu spesifik yang sedang hangat diperbincangkan.

Salah satu isu yang menjadi sorotan publik di Indonesia adalah Badan Pengelola Investasi (BPI) Danantara, lembaga yang bertugas mengelola kekayaan negara untuk mendanai investasi strategis. Sejak pendiriannya, Danantara telah memicu beragam reaksi dari masyarakat, mulai dari apresiasi hingga tanggapan kritis terkait transparansi dan efektivitasnya. Diskusi tentang Danantara di platform X mencerminkan dinamika opini publik yang kompleks dan beragam. Namun, analisis sentimen terhadap isu ini menghadapi tantangan besar, terutama karena bahasa Indonesia memiliki struktur dan variasi yang unik, serta data dari media sosial sering kali mengandung *noise* seperti emoji, slang, dan kesalahan penulisan.

Metode analisis sentimen tradisional, seperti *Naive Bayes* dan *SVM*, sering kali gagal menangkap konteks kompleks dalam teks bahasa Indonesia [3], [1]. Hal ini disebabkan oleh keterbatasan metode tersebut dalam memahami makna kata berdasarkan konteks penggunaannya. Selain itu, penelitian sebelumnya yang menggunakan model berbasis *transformers* untuk analisis sentimen masih terbatas, terutama dalam konteks isu-isu strategis seperti Danantara. Oleh karena itu, diperlukan pendekatan yang lebih canggih untuk meningkatkan akurasi analisis sentimen dan menggali wawasan yang lebih mendalam tentang opini publik.

Penelitian ini menggunakan algoritma IndoBERT, sebuah model berbasis *transformers* yang telah dioptimalkan untuk bahasa Indonesia [4]. Dengan kemampuannya memproses teks secara *bidirectional*, IndoBERT mampu menangkap konteks kata berdasarkan informasi sebelum dan sesudahnya [5], [6]. Hal ini menjadikan IndoBERT lebih unggul dibandingkan metode analisis sentimen tradisional, terutama dalam menangani teks yang kompleks. Penelitian ini dimaksudkan

untuk menganalisis serta menilai opini masyarakat mengenai Badan Pengelola Investasi (BPI) Danantara menggunakan algoritma IndoBERT, dengan memanfaatkan data besar dari media sosial X. Penelitian ini juga bertujuan untuk memetakan pola sentimen yang dominan (positif, negatif, atau netral) secara akurat dan sistematis, serta juga menyajikan hasil kuantitatif yang bisa mendukung para pembuat kebijakan dalam memperoleh pemahaman yang lebih komprehensif mengenai persepsi masyarakat.

1.2 Rumusan Masalah

Merujuk pada pembahasan latar belakang, berikut ini adalah rumusan masalah penelitian yang diangkat:

1. Pola sentimen mana yang paling dominan (positif, negatif, atau netral) dalam opini masyarakat terhadap Danantara?
2. Sejauh mana keakuratan analisis sentimen menggunakan IndoBERT dalam menangkap persepsi masyarakat terhadap Danantara?
3. Bagaimana penerapan algoritma IndoBERT dalam menganalisis sentimen publik terhadap Badan Pengelola Investasi (BPI) Danantara berdasarkan data media sosial X?

1.3 Tujuan dan Manfaat

1.3.1 Tujuan Penelitian

Capaian sasaran yang diinginkan melalui akhir studi ini mencakup:

1. Menemukan pola sentimen dominan yang muncul dalam diskusi terkait isu Danantara dan memberikan interpretasi terhadap pola tersebut.
2. Mengevaluasi kemampuan algoritma IndoBERT dalam menghasilkan analisis sentimen yang akurat terhadap opini publik berbasis bahasa Indonesia.
3. Menganalisis dinamika opini publik terkait isu Danantara melalui analisis sentimen pada platform X dengan menggunakan data kuantitatif.

1.3.2 Manfaat Penelitian

Proyek diharapkan dapat memberikan kontribusi mencakup:

1. Berperan guna memperkaya perkembangan wawasan keilmuan di bidang analisis sentimen menggunakan algoritma *machine learning*, terutama IndoBERT pada data berbahasa Indonesia.
2. Memberikan gambaran yang lebih terstruktur mengenai dinamika opini publik terkait isu Danantara melalui analisis sentimen berbasis media sosial
3. Menjadi acuan bagi peneliti dan praktisi dalam pengembangan penelitian terkait analisis sentimen menggunakan data besar di platform media sosial
4. Mendorong penggunaan teknologi kecerdasan buatan untuk memahami dinamika sosial dalam isu-isu nasional.

1.4 Batasan Masalah

Penelitian ini menetapkan sejumlah batasan yang harus dicermati supaya cakupan penelitian lebih terarah dan mudah dipahami. Beberapa batasan tersebut antara lain:

1. Studi ini hanya meneliti hal-hal yang diunggah orang di satu situs media sosial bernama platform X. Studi ini tidak mencakup unggahan dari situs lain seperti Instagram, Facebook, Tiktok, atau YouTube.
2. Studi ini hanya memeriksa unggahan yang dibuat antara 1 Januari 2025 hingga 22 April 2025. Unggahan dari sebelum atau sesudah tanggal tersebut tidak diteliti.
3. Penelitian ini hanya menggunakan algoritma IndoBERT sebagai metode utama untuk menganalisis sentimen, tanpa membandingkannya dengan algoritma atau metode lain.
4. Penelitian ini terbatas pada analisis sentimen dalam bahasa Indonesia, sehingga data dalam bahasa lain tidak akan dianalisis.
5. Penelitian ini hanya berfokus pada klasifikasi sentimen (positif, negatif, netral) terkait isu Danantara. Aspek lain seperti analisis emosi spesifik (marah, senang, sedih) atau analisis topik tidak dibahas.
6. Dataset yang digunakan berasal dari unggahan publik di platform X. Data dari akun privat, komentar tersembunyi, atau sumber lain seperti artikel berita tidak termasuk dalam penelitian ini.
7. Fokus penelitian adalah pada isu Danantara, sehingga hasil analisis tidak dapat digeneralisasikan untuk isu-isu lain.

1.5 Keterbaruan Penelitian

Pada bagian ini telah dirangkum keterbaruan penelitian selaras dengan berkas literatur yang telah ada sebelumnya mendiskusikan penerapan algoritma BERT, diantaranya:

1. Dari penelitian yang dilakukan oleh Geni Lenggo, Yulianti, Sensuse Indra (2023) dalam artikel mengangkat “Analisis Sentimen *Tweet* Sebelum Pemilu 2024 di Indonesia Menggunakan Model Bahasa IndoBERT” menyampaikan model baseline dengan BERT ada disetujui oleh *TextBlob* (48,5% dan 46,49%), *Naive Bayes* (3,5%), serta *Support Vector Machine* (13,49%). Hasil terbaik IndoBERT 83,5% [7].
2. Pada karya yang diusulkan oleh Saraswati, Vania dan Yudhistira dengan Putra Pandu Adikara berjudul “Analisis Sentimen terhadap perundungan Siber pada Twitter menggunakan Algoritma *Bidirectional Encoder Representation from transformers* (BERT)” diperoleh hasil dari test dan validasi berbahasa Indonesia, di mana BERT menunjukkan presisi 81%, recall 80.67%, serta F1-score 80.67% [8].
3. Penjelasan selanjutnya disampaikan oleh Kurniawan, Aldino dan Isnain dalam artikel tahun 2022 berjudul “Sentimen Analisis Terhadap Kebijakan Penyelenggara Sistem Elektronik (PSE) menggunakan *Algoritma Bidirectional Encoder Representations from transformers* (BERT)” yang melaporkan akurasi 69,55% dan 55% pada 2 periode waktu berbeda dengan *hyperparameter batch size 16, epoch 5*. Menariknya, akurasi pada data seimbang (62%) lebih tinggi meskipun dalam jumlah lebih sedikit dibandingkan data tidak seimbang [9].
4. Adine Nayla, Casi Setianingsih, dan Burhanuddin Dirgantoro (2023) pada penelitian berjudul “Deteksi *Hate speech* pada Twitter menggunakan Algoritma BERT” memperoleh hasil pengujian akurasi 78,65%, presisi 78,90%, *recall* 78,69%, dan *F1-score* 78,77% untuk klasifikasi *hate speech* [10].