

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi telah meningkatkan jumlah data digital yang dihasilkan dari berbagai platform media *online*, sehingga menimbulkan tantangan dalam bidang informatika untuk memproses data yang bersifat *multimodal*, yaitu data yang menggabungkan teks dan gambar secara bersamaan.

Pendekatan *multimodal learning* merupakan metode pembelajaran mesin yang mengintegrasikan lebih dari satu jenis data, seperti teks dan visual, agar sistem kecerdasan buatan mampu memahami konteks informasi secara lebih menyeluruh dan akurat [1], [2]. Pendekatan ini banyak diterapkan dalam berbagai bidang, termasuk analisis media sosial, klasifikasi berita, dan deteksi informasi palsu karena mampu memanfaatkan hubungan semantik antar-modalitas dalam suatu informasi digital.

Fenomena penyebaran berita palsu (*fake news*) di Indonesia menjadi salah satu permasalahan yang mendapat perhatian dalam ruang digital. Data Kementerian Komunikasi dan Digital (KomDigi) mencatat lebih dari 12.000 konten hoaks teridentifikasi sejak tahun 2018 hingga 2024 [3], [4]. Berita palsu umumnya memanfaatkan kombinasi teks dan gambar untuk membangun persepsi tertentu terhadap informasi yang disampaikan [5]. Penggunaan gambar yang tidak relevan, diambil dari konteks berbeda, atau tidak sesuai dengan isi teks sering digunakan untuk memperkuat klaim tertentu sehingga memengaruhi persepsi pembaca terhadap suatu informasi. Kondisi tersebut menyebabkan masyarakat menghadapi kesulitan dalam membedakan informasi kredibel dan tidak kredibel apabila penilaian hanya dilakukan berdasarkan teks.

Kredibilitas berita merupakan tingkat kepercayaan terhadap informasi yang dipengaruhi oleh konsistensi, akurasi, dan kesesuaian elemen informasi yang disajikan [6]. Dalam lingkungan digital, penilaian terhadap kredibilitas

informasi tidak hanya bergantung pada isi teks, tetapi juga pada kesesuaian antara teks dan representasi visual yang ditampilkan [7], [8]. Ketidaksesuaian antara gambar dan isi berita dapat memengaruhi persepsi pembaca terhadap validitas informasi yang disampaikan [9]. Oleh karena itu, hubungan semantik antara teks dan gambar dapat dijadikan salah satu indikator penting dalam proses evaluasi kredibilitas berita digital.

Klasifikasi kredibilitas berita berbasis *Multimodal* memiliki kompleksitas yang lebih tinggi dibandingkan pendekatan unimodal karena sistem harus mampu memahami hubungan semantik antara teks dan gambar secara bersamaan. Informasi visual dan tekstual dalam berita digital sering kali saling melengkapi, sehingga diperlukan mekanisme representasi yang mampu menangkap keterkaitan antar-modalitas secara efektif.

Dalam pendekatan komputasional, konsistensi antara teks dan gambar dapat direpresentasikan menggunakan perhitungan *cosine similarity* antara *embedding* teks yang dihasilkan *Indonesian Bidirectional Encoder Representations from Transformers (IndoBERT)* dan *embedding* visual yang dihasilkan *Contrastive Language Image Pretraining (CLIP)*. Nilai *similarity* yang tinggi menunjukkan tingkat kesesuaian semantik antar-modalitas, sedangkan nilai yang rendah dapat mengindikasikan potensi inkonsistensi informasi yang berkaitan dengan rendahnya kredibilitas berita. Oleh karena itu, diperlukan integrasi model bahasa dan model vision-language agar representasi *multimodal* dapat dianalisis secara terukur.

Selain itu, pendekatan *multimodal* juga memiliki tantangan dalam memahami konteks bahasa Indonesia sekaligus memetakan hubungan visual dan tekstual ke dalam ruang *embedding* yang sama. Pendekatan berbasis teks saja belum mampu menganalisis hubungan antara isi teks dan gambar yang sering dimanfaatkan dalam penyebaran berita palsu [5]. Oleh karena itu, diperlukan pendekatan *multimodal* yang dapat memproses kedua modalitas secara bersamaan agar analisis kredibilitas berita dapat dilakukan secara lebih komprehensif.

Penerapan pendekatan *multimodal* dalam penelitian terdahulu menunjukkan kontribusi yang signifikan terhadap peningkatan performa klasifikasi informasi. Zhou et al. (2022) melaporkan bahwa integrasi model bahasa dengan *CLIP* mampu meningkatkan akurasi deteksi berita palsu sebesar 6,8% dibandingkan pendekatan unimodal [1]. Selain itu, Jiang dan Wang (2025) menunjukkan bahwa pendekatan berbasis vision-language model mampu mencapai akurasi sebesar 95,6% dengan *F1-Score* sebesar 94,8% pada tugas deteksi berita palsu *multimodal* [2]. Hasil penelitian tersebut menunjukkan bahwa hubungan antara teks dan gambar memiliki peran penting dalam memahami konteks informasi secara lebih komprehensif dibandingkan penggunaan satu jenis data saja.

Namun, dalam konteks lokal Indonesia, beberapa penelitian terkait klasifikasi berita hoaks masih didominasi oleh pendekatan berbasis teks (unimodal). Penelitian mengenai klasifikasi hoaks berita bencana alam umumnya hanya memanfaatkan fitur linguistik dari isi berita tanpa mempertimbangkan informasi visual yang menyertainya [10], [11]. Salah satu penelitian menggunakan representasi *IndoBERT* dan algoritma *XGBoost* untuk mendeteksi berita hoaks bencana alam berbahasa Indonesia serta menunjukkan hasil akurasi sebesar 99,11% dengan *F1-Score* sebesar 0,99 dalam mengenali pola linguistik berita hoaks. Penelitian tersebut menunjukkan bahwa *IndoBERT* efektif dalam memahami konteks semantik teks berbahasa Indonesia melalui representasi kontekstual yang dihasilkan dari model berbasis *Transformer*. Meskipun demikian, pendekatan yang digunakan masih berfokus pada modalitas teks sehingga belum mempertimbangkan hubungan antara isi teks dan gambar dalam berita online.

Berdasarkan uraian tersebut, masih terdapat celah penelitian dalam pengembangan sistem klasifikasi kredibilitas berita berbasis *Multimodal* di Indonesia yang secara eksplisit memanfaatkan konsistensi informasi antara teks dan gambar sebagai indikator penilaian kredibilitas. Penelitian terkait integrasi *IndoBERT* dan *CLIP* untuk analisis kredibilitas berita online

berbahasa Indonesia juga masih terbatas, khususnya pada pengukuran keselarasan semantik antar-modalitas menggunakan pendekatan *Embedding*.

Untuk mendukung pendekatan tersebut, penelitian ini mengusulkan integrasi *IndoBERT* dan *CLIP* dalam klasifikasi kredibilitas berita online berbasis *Multimodal*. *IndoBERT* merupakan model bahasa berbasis *Transformer* yang dilatih menggunakan korpus bahasa Indonesia dalam jumlah besar sehingga mampu memahami konteks linguistik, struktur sintaksis, serta makna semantik bahasa Indonesia secara lebih baik dibandingkan model bahasa umum [12]. *IndoBERT* juga memiliki kemampuan unggul dalam menangkap hubungan kontekstual antar kata melalui mekanisme *bidirectional contextual representation*, sehingga efektif digunakan dalam berbagai tugas Natural Language Processing (NLP) berbahasa Indonesia, termasuk klasifikasi teks dan analisis sentimen.

Sementara itu, *CLIP* merupakan model *vision-language* yang mampu memetakan representasi teks dan gambar ke dalam ruang *Embedding* yang sama sehingga hubungan semantik antar-modalitas dapat dianalisis secara lebih efektif [1], [7]. Integrasi kedua model tersebut menjadi kebaruan penelitian karena memungkinkan pengukuran keselarasan makna antara teks dan gambar dalam berita online menggunakan pendekatan *Multimodal* berbasis *Embedding*.

Dalam penelitian ini, data yang digunakan berupa judul berita dan gambar yang diperoleh dari hasil *scraping* berbagai portal berita serta situs pemeriksa fakta. Judul berita digunakan sebagai representasi teks karena mampu merangkum informasi utama yang disampaikan dan tersedia secara konsisten pada seluruh sumber data. Selanjutnya, pengukuran keselarasan antar-modalitas dilakukan menggunakan *Cosine Similarity* sebagai representasi kuantitatif dari konsistensi informasi antara teks dan gambar. Nilai *similarity* yang dihasilkan diharapkan mampu membantu sistem dalam mengidentifikasi tingkat kredibilitas berita berdasarkan kesesuaian semantik antar-modalitas. Oleh karena itu, penelitian ini diharapkan mampu menghasilkan sistem klasifikasi

kredibilitas berita yang lebih objektif, terukur, dan lebih efektif dibandingkan pendekatan berbasis teks saja.

1.2 Rumusan Masalah

1. Bagaimana mengintegrasikan model *IndoBERT* dan *CLIP* dalam mengklasifikasikan kredibilitas berita *online* berbasis *Multimodal*?
2. Bagaimana kinerja model *Multimodal* hasil integrasi *IndoBERT* dan *CLIP* dalam mengklasifikasikan kredibilitas berita *online*?

1.3 Tujuan Penelitian

1. Mengembangkan sistem klasifikasi kredibilitas berita online berbasis *Multimodal* dengan mengombinasikan representasi teks dari *IndoBERT* dan representasi visual dari *CLIP*.
2. Mengevaluasi kinerja sistem klasifikasi menggunakan metrik evaluasi klasifikasi, yaitu *Accuracy*, *Precision*, *Recall*, dan *F1-Score*, untuk menilai efektivitas pendekatan yang diusulkan.

1.4 Batasan Masalah

1. Penelitian ini menggunakan *Dataset* sebanyak 1.344 berita *online* berbahasa Indonesia yang berasal dari portal berita nasional dan situs pemeriksa fakta seperti Kompas, CNN Indonesia, Detik, TurnBackHoax.id, Kompas Cek Fakta, serta Liputan6 Cek Fakta.
2. Data yang digunakan dibatasi hanya pada berita yang memiliki pasangan teks dan gambar agar sesuai dengan kebutuhan analisis *Multimodal*, sehingga konteks penyebaran berita digital di Indonesia dapat direpresentasikan secara lebih relevan.
3. Ruang lingkup penelitian dibatasi pada berita bertema bencana alam, mencakup topik seperti gempa bumi, banjir, tanah longsor, letusan gunung berapi, dan cuaca ekstrem, guna menjaga konsistensi domain serta relevansi isu terhadap konteks sosial dan lingkungan nasional.

4. Pengumpulan data dilakukan pada periode 2024–2025 untuk menyesuaikan dengan tren penyebaran berita terkini serta memastikan relevansi temporal data yang digunakan dalam analisis kredibilitas.
5. Model yang digunakan terdiri dari *IndoBERT* sebagai text encoder dan *CLIP* sebagai *image encoder* yang digabungkan melalui *Fusion* layer untuk menghasilkan representasi *Multimodal* teks–gambar dalam analisis tingkat kredibilitas berita.
6. Penelitian ini menerapkan metodologi *Cross Industry Standard Process for Data Mining (CRISP-DM)* yang meliputi tahapan *Business Understanding, Data Understanding, Data Preparation, Modeling, dan Evaluation*. Penelitian dibatasi hanya sampai tahap evaluasi tanpa mencakup tahap *Deployment*.

1.5 Manfaat

Penelitian tugas akhir ini diharapkan memberikan beberapa manfaat, antara lain:

1.5.1 Manfaat Akademik

1. Memberikan kontribusi ilmiah dalam pengembangan ilmu komputer, khususnya pada bidang *Artificial Intelligence* berbasis *Multimodal* yang mengintegrasikan Natural Language Processing (NLP) dan Computer Vision (CV) dalam konteks analisis kredibilitas berita berbahasa Indonesia.
2. Memperkaya literatur akademik mengenai pemanfaatan model *Transformer*, khususnya *IndoBERT* dan *CLIP*, dalam mengukur keselarasan semantik antara teks dan gambar sebagai indikator kredibilitas informasi digital.

1.5.2 Manfaat bagi Masyarakat

1. Mendukung terciptanya ekosistem informasi digital yang lebih terpercaya melalui penyediaan mekanisme penilaian kredibilitas berita yang berbasis data dan keselarasan makna antara teks dan gambar, bukan semata penilaian subjektif.
2. Membantu masyarakat dalam meningkatkan literasi digital, khususnya

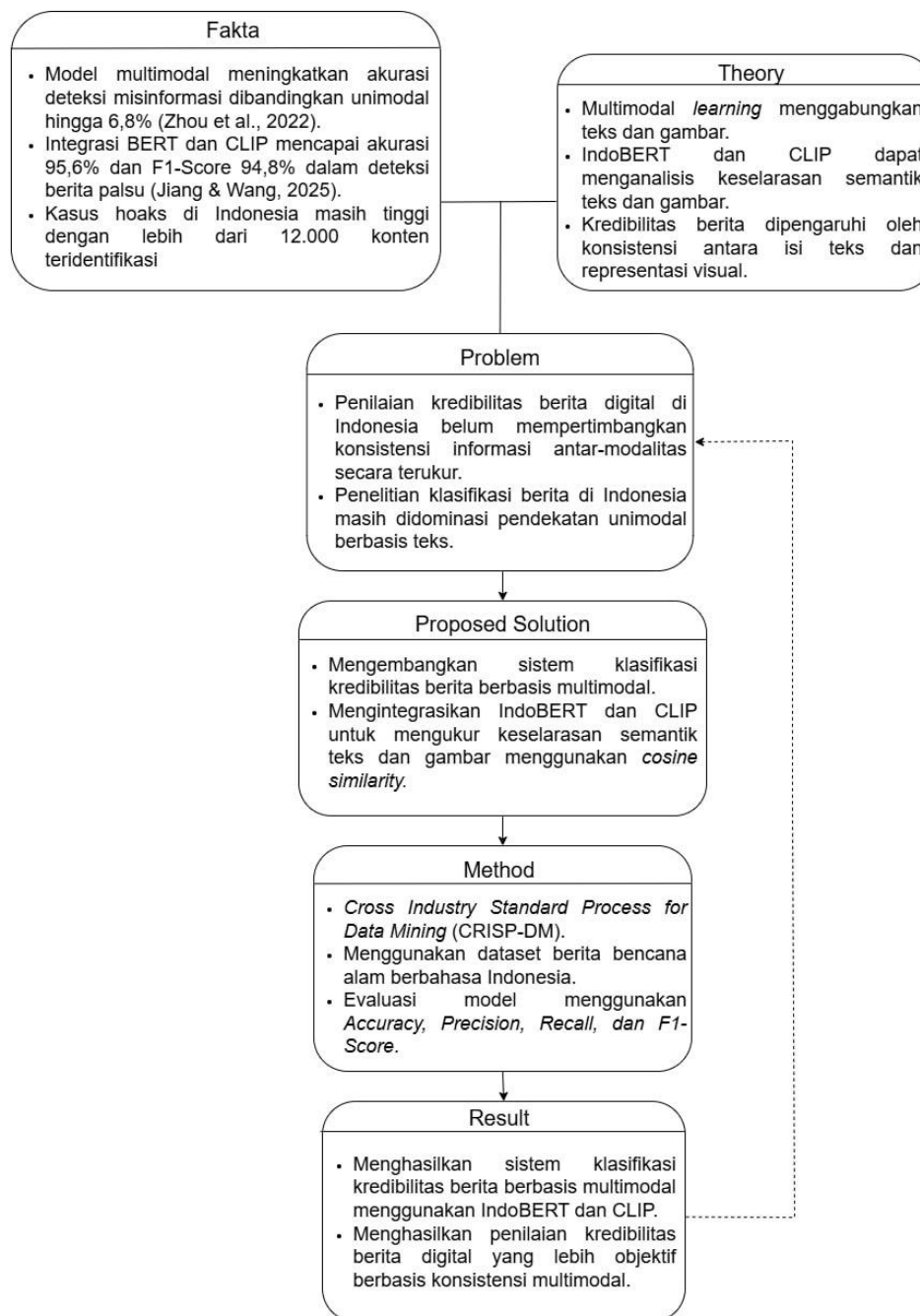
dalam menyaring dan mengevaluasi berita bencana alam di media online, sehingga dapat mengurangi dampak negatif penyebaran tidak kredibel dan misinformasi.

1.5.3 Manfaat bagi Peneliti

1. Menjadi referensi empiris bagi peneliti selanjutnya yang ingin mengembangkan sistem deteksi tidak kredibel, klasifikasi kredibilitas berita, atau analisis misinformasi berbasis *Multimodal Learning*.
2. Memberikan wawasan teknis mengenai implementasi penggabungan *Embedding* teks dan visual, perhitungan *Cosine Similarity*, serta evaluasi performa model *Multimodal* dalam menangani data berita online yang kompleks dan tidak terstruktur.



1.6 Kerangka Berpikir



Gambar 1.1 Kerangka Berpikir

Gambar 1.1 menggambarkan kerangka berpikir penelitian yang diawali dari fakta bahwa pendekatan *Multimodal* terbukti mampu meningkatkan performa deteksi misinformasi dibandingkan pendekatan *unimodal*. Penelitian Zhou et al. (2022) menunjukkan bahwa model *Multimodal* dapat meningkatkan akurasi deteksi misinformasi hingga 6,8%, sedangkan Jiang dan Wang (2025) melaporkan bahwa integrasi *BERT* dan *CLIP* mampu mencapai akurasi sebesar

95,6% dan *F1-Score* sebesar 94,8% dalam deteksi berita palsu. Selain itu, tingginya jumlah kasus hoaks di Indonesia dengan lebih dari 12.000 konten teridentifikasi menunjukkan bahwa permasalahan kredibilitas informasi digital masih menjadi isu yang signifikan. Kerangka berpikir ini juga didasarkan pada teori *Multimodal Learning* yang menyatakan bahwa pemahaman informasi dapat dilakukan secara lebih komprehensif melalui penggabungan modalitas teks dan gambar. Dalam konteks tersebut, *IndoBERT* dan *CLIP* digunakan untuk menganalisis keselarasan semantik antara teks dan gambar, sedangkan kredibilitas berita dipengaruhi oleh konsistensi antara isi teks dan representasi visual yang ditampilkan.

Berdasarkan fakta dan teori tersebut, penelitian ini mengidentifikasi permasalahan bahwa penilaian kredibilitas berita digital di Indonesia belum mempertimbangkan konsistensi informasi antar-modalitas secara terukur. Selain itu, penelitian klasifikasi berita di Indonesia masih didominasi pendekatan *unimodal* berbasis teks sehingga hubungan semantik antara teks dan gambar belum dianalisis secara optimal. Untuk mengatasi permasalahan tersebut, penelitian ini mengusulkan pengembangan sistem klasifikasi kredibilitas berita berbasis *Multimodal* dengan mengintegrasikan *IndoBERT* dan *CLIP*. Integrasi kedua model dilakukan untuk mengukur keselarasan semantik antara teks dan gambar menggunakan *Cosine Similarity* sebagai representasi kuantitatif konsistensi informasi *Multimodal*. Metode penelitian menggunakan pendekatan *Cross Industry Standard Process for Data Mining (CRISP-DM)* dengan *Dataset* berita bencana alam berbahasa Indonesia. Evaluasi model dilakukan menggunakan metrik *Accuracy*, *Precision*, *Recall*, dan *F1-Score* untuk mengukur performa klasifikasi yang dihasilkan. Hasil yang diharapkan dari penelitian ini adalah terbentuknya sistem klasifikasi kredibilitas berita berbasis *Multimodal* menggunakan *IndoBERT* dan *CLIP* yang mampu menghasilkan penilaian kredibilitas berita digital secara lebih objektif dan terukur berdasarkan konsistensi antara teks dan gambar.