

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di tengah arus digitalisasi yang semakin masif membuat berita finansial daring menjadi sumber informasi utama bagi investor dan pelaku pasar modal. Setiap perilsan berita, analisis ekonomi hingga opini publik di media dapat memantik reaksi cepat yang dapat dilihat dari fluktuasi harga saham dan indeks sektoral. Di Indonesia, sektor keuangan direpresentasikan oleh Indeks Sektoral Keuangan (IDXFINANCE) di Bursa Efek Indonesia (BEI). Indeks tersebut memegang peran penting sebagai fondasi stabilitas ekonomi nasional. Namun, sektor ini sangat dipengaruhi oleh persepsi publik sehingga menjadi rentan terhadap berbagai narasi di ruang digital, baik yang bersumber dari faktor fundamental seperti laporan laba dan dividen, maupun faktor non-fundamental seperti sentimen investor dan berita makroekonomi yang viral.

Dalam era *Big Data*, ribuan berita finansial diterbitkan setiap hari dari berbagai portal nasional dan internasional. Volume data yang masif ini menjadi tantangan besar bagi investor, analis, dan regulator untuk menyaring isu-isu strategis dan narasi risiko secara cepat dan akurat. Pendekatan tradisional, seperti analisis manual atau metode statistik konvensional, kini mulai kehilangan efektivitasnya karena keterbatasan skalabilitas dan kecepatan pemrosesan. Hal ini mendorong lembaga keuangan untuk bertransformasi menuju pengawasan dan pengambilan keputusan berbasis analitik data besar. Sebagai bukti nyata, Otoritas Jasa Keuangan (OJK) telah meluncurkan aplikasi *Suptech Integrated Analytics Data* (OSIDA PMDK) sebagai bagian dari strategi transformasi pengawasan pasar modal berbasis teknologi dan data terintegrasi [1].

Salah satu pendekatan yang berkembang pesat dalam menjawab tantangan tersebut adalah pemodelan topik (*topic modeling*), yaitu metode yang berupaya menemukan pola atau struktur tematik tersembunyi di balik kumpulan teks berskala besar salah satu nya *Latent Dirichlet Allocation* (LDA). Dalam berbagai

penelitian, LDA telah terbukti efektif untuk eksplorasi topik berita finansial, termasuk dalam konteks Indonesia [2]. Namun, seiring berkembangnya teknologi *Natural Language Processing* (NLP), model berbasis *transformer embeddings* seperti BERTopic memiliki kemampuan yang lebih kontekstual. Berbeda dari LDA yang bergantung pada distribusi kata, BERTopic menggabungkan representasi semantik dari model *language transformer* (misalnya BERT atau *Sentence-BERT*) dengan teknik *clustering* seperti HDBSCAN. Hasilnya, model ini mampu mengelompokkan berita dengan mempertimbangkan makna kontekstual, bukan sekadar kemunculan kata yang sama [3]. Sejumlah studi terkini menunjukkan bahwa BERTopic menghasilkan koherensi topik yang lebih tinggi dibandingkan LDA terutama pada teks non-terstruktur seperti berita ekonomi dan media sosial [4].

Melihat dinamika tersebut, penelitian ini berupaya melakukan analisis komparatif antara LDA dan BERTopic dalam mengidentifikasi topik-topik utama dari berita pasar finansial Indonesia. Lebih lanjut, penelitian ini juga mengintegrasikan analisis sentimen untuk menelusuri bagaimana narasi positif maupun negatif di media dapat mempengaruhi persepsi risiko pasar dan pergerakan Indeks Sektoral Keuangan (IDXFINANCE). Pendekatan ini diharapkan mampu menemukan hubungan antara narasi publik dan performa sektor keuangan nasional. Selain itu, pendekatan ini diharapkan turut serta memberikan kontribusi pada pengembangan *early warning system* berbasis analitik-teks di pasar modal Indonesia.

1.2 Rumusan Masalah

Dari latar belakang tersebut diperlukan pendekatan untuk mencapai tujuan dari penelitian ini. Berikut rumusan masalah yang dapat diambil:

1. Bagaimana perbandingan performa LDA dan BERTopic dalam mengidentifikasi topik utama dari berita pasar finansial Indonesia?

2. Bagaimana integrasi analisis sentimen berita dapat memperkuat interpretasi hasil pemodelan topik untuk mendeteksi narasi risiko dan peluang pasar?
3. Sejauh mana narasi dan sentimen yang diidentifikasi dari berita finansial memiliki korelasi dengan pergerakan Indeks Sektoral Keuangan (IDXFİNANCE) di Bursa Efek Indonesia?

1.3 Batasan Masalah

Penelitian ini memiliki beberapa batasan untuk menjaga fokus dan ruang lingkup pembahasan, yaitu:

1. Objek Penelitian

Penelitian ini berfokus pada berita finansial daring yang berkaitan dengan pasar modal Indonesia, khususnya yang memiliki keterkaitan dengan sektor keuangan. Berita diambil dari portal berita nasional yang kredibel yaitu Bisnis Indonesia, CNBC Indonesia.

2. Periode Data

Pergerakan berita dan data indeks sektoral yang dianalisis dibatasi pada periode tertentu agar analisis dapat dibandingkan dengan pergerakan historis IDXFİNANCE dalam periode yang sama. Periode waktu dalam penelitian ini yaitu bulan Mei hingga September 2025.

3. Model Pemodelan Topik

Penelitian hanya membandingkan dua metode utama, yaitu *Latent Dirichlet Allocation* (LDA) dan BERTopic, tanpa memasukkan model lain seperti NMF atau Top2Vec. Perbandingan difokuskan pada koherensi topik, interpretabilitas, dan relevansi semantik hasil.

4. Analisis Sentimen

Analisis sentimen dilakukan menggunakan model berbasis bahasa Indonesia (seperti *IndoBERT Sentiment*) dengan pendekatan klasifikasi polaritas seperti positif, negatif, netral.

5. Korelasi dengan IDXFINANCE

Korelasi yang dikaji terbatas pada hubungan statistik antara dinamika sentimen/narasi berita dan pergerakan indeks sektoral keuangan (IDXFINANCE). Penelitian tidak melakukan analisis kausalitas atau prediksi harga saham secara individual.

6. Konteks Bahasa dan Data

Data teks yang digunakan sepenuhnya dalam bahasa Indonesia. Berita berbahasa asing atau terjemahan otomatis tidak disertakan untuk menjaga konsistensi hasil *topic modeling* dan *sentiment analysis*.

7. Aspek Ekonomi dan Kebijakan

Penelitian ini tidak membahas faktor-faktor ekonomi makro secara mendalam seperti suku bunga, inflasi, atau nilai tukar, kecuali sebagai konteks tambahan dalam interpretasi hasil narasi berita.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Melakukan analisis komparatif antara metode *topic modeling* berbasis LDA dan BERTopic dalam mengidentifikasi tema utama yang muncul dari berita pasar finansial Indonesia.
2. Mengintegrasikan analisis sentimen ke dalam hasil pemodelan topik guna memahami persepsi pasar dan arah narasi risiko yang terbentuk di ruang publik.
3. Menguji korelasi antara hasil analisis topik dan sentimen berita finansial dengan pergerakan indeks sektoral IDXFINANCE, sebagai representasi kondisi pasar keuangan nasional.
4. Memberikan kontribusi empiris dan metodologis bagi pengembangan sistem pemantauan berbasis *Big Data Analytics* di sektor keuangan Indonesia, yang dapat mendukung kebijakan pengawasan pasar modal berbasis narasi dan sentimen digital.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut.

1. Bagi Investor:

- a. Memberikan bantuan analitis dalam memahami berita keuangan dan kondisi pasar.
- b. Membantu pengambilan keputusan investasi yang didasari oleh data berita.

2. Bagi Akademisi dan Peneliti:

- a. Menambah referensi kajian mengenai penerapan NLP dalam bidang keuangan dan pasar modal.
- b. Menjadi model awal untuk riset lanjutan mengenai sistem rekomendasi berbasis data berita.

3. Bagi Perusahaan atau Startup Fintech:

- a. Menjadi dasar pengembangan aplikasi rekomendasi investasi berbasis AI yang bisa diadopsi secara komersial.
- b. Mendukung inklusi keuangan dengan menyediakan solusi teknologi yang membantu investor.

4. Bagi Pemerintah atau Otoritas Keuangan (OJK, IDX):

- a. Mendukung program literasi dan inklusi keuangan melalui pemanfaatan teknologi untuk dasar edukasi dan rekomendasi investasi.

1.6 Sistematika Penulisan

1. BAB I PENDAHULUAN

Berisi latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan.

2. BAB II TINJAUAN PUSTAKA

Menyajikan landasan teori yang relevan, penelitian terdahulu, serta kerangka berpikir untuk mendukung analisis yang akan dilakukan.

3. BAB III METODE PENELITIAN

Menjelaskan metode yang digunakan dalam penelitian, mulai dari pengumpulan data, *preprocessing*, *topic modeling*, sentimen analisis, dan analisis korelasi IDXFINANCE.

4. BAB IV HASIL DAN PEMBAHASAN

Memaparkan analisis komparatif, integrasi analisis sentimen, dan korelasi indeks sektoral IDXFINANCE.

5. BAB V PENUTUP

Menyajikan kesimpulan dari hasil penelitian dan memberikan saran untuk pengembangan lebih lanjut baik dari sisi teknis maupun fungsionalitas sistem.

