

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi dan komunikasi telah membawa perubahan besar dalam cara masyarakat memperoleh, menyebarkan, dan memproses informasi, termasuk dalam hal investasi. Salah satu bentuk investasi yang mengalami transformasi digital adalah emas digital, di mana emas tidak hanya menjadi instrumen investasi konvensional, tetapi juga telah merambah ke bentuk digital melalui platform-platform finansial berbasis aplikasi (Ginting et al., 2025). Kemajuan teknologi yang pesat juga mengubah cara orang berinvestasi, mendorong hampir semua perusahaan untuk berlomba-lomba menciptakan platform digital agar dapat terus memberikan pelayanan kepada konsumen, terutama dalam hal platform investasi (Maulana et al., 2024).

Pada era digital saat ini, investasi telah menjadi bagian penting dari bisnis dan perekonomian. Data dari Kustodian Sentral Efek Indonesia (KSEI) menunjukkan bahwa jumlah investor pasar modal Indonesia mencapai 12,47 juta orang pada bulan Februari 2024, mengalami kenaikan sebesar 1,20% dari bulan Januari 2024 (Annisa et al., 2024). Transformasi digital telah mempermudah akses ke investasi melalui berbagai aplikasi investasi yang kini dapat diakses melalui perangkat smartphone Android maupun iPhone (Annisa et al., 2024). Hal ini menunjukkan bahwa minat masyarakat terhadap berbagai instrumen investasi digital, termasuk emas digital, terus mengalami pertumbuhan yang signifikan.

Seiring meningkatnya tren emas digital, persepsi dan opini publik terhadap produk ini menjadi topik yang menarik untuk diteliti. Opini masyarakat yang disampaikan melalui media sosial dapat memberikan gambaran nyata tentang penerimaan publik terhadap inovasi finansial tersebut (Ginting et al., 2025). Di era digital, opini masyarakat terkait perusahaan atau isu tertentu kini mudah ditemukan di media sosial, termasuk Twitter, yang menjadi platform penting dalam interaksi sosial dan analisis sentimen (Norlaila et al., 2024). Memilih aplikasi investasi online juga membutuhkan informasi yang dikumpulkan dari pengalaman pengguna atau ulasan dari aplikasi investasi tersebut, dan masyarakat dapat memanfaatkan

sentimen ini untuk memahami informasi yang terkait dengan aplikasi investasi (Annisa et al., 2024).

Untuk memahami dan mengolah opini publik yang tersebar luas tersebut, dibutuhkan suatu pendekatan analisis data yang sistematis. Analisis sentimen merupakan proses menganalisis teks untuk menentukan sentimen yang diungkapkan, seperti positif, negatif, atau netral, yang dapat digunakan untuk berbagai tujuan seperti memahami opini publik, mengukur kepuasan pelanggan, dan mendeteksi penipuan (Maulana et al., 2024). Sentimen analisis atau opinion mining merupakan cabang dari Natural Language Processing (NLP) yang bertujuan untuk mengidentifikasi dan mengklasifikasikan opini pengguna terhadap suatu entitas ke dalam kategori positif, negatif, atau netral (Ginting et al., 2025). Teknik ini memungkinkan peneliti untuk menangkap sikap emosional pengguna terhadap topik tertentu dari data teks yang tidak terstruktur, seperti tweet, komentar, atau ulasan (Ginting et al., 2025).

Penelitian terdahulu yang secara khusus membahas sentimen masyarakat terhadap emas digital menggunakan Naïve Bayes telah dilakukan oleh (Ginting et al., 2025). Penelitian tersebut mengumpulkan data dari Twitter selama periode tiga bulan dengan menggunakan Twitter API, kemudian melakukan proses pra-pemrosesan teks dan menerapkan model Naïve Bayes untuk mengklasifikasikan opini menjadi sentimen positif, negatif, dan netral. Dari total 5.000 tweet yang dikumpulkan dan dianalisis, diperoleh hasil klasifikasi sebagai berikut: 52% tweet bersentimen positif, 30% bersentimen negatif, dan 18% bersentimen netral. Temuan ini menunjukkan bahwa mayoritas masyarakat memiliki pandangan yang positif terhadap emas digital, terutama terkait dengan kemudahan akses, keamanan transaksi, dan potensi investasi. Namun sentimen negatif yang muncul masih cukup signifikan, mengindikasikan adanya keraguan dan kekhawatiran pengguna terhadap layanan atau platform emas digital tertentu, sehingga hal ini menjadi catatan penting bagi penyedia layanan untuk meningkatkan kualitas, transparansi, dan literasi digital kepada konsumen (Ginting et al., 2025).

Selain itu, penerapan Naïve Bayes dalam analisis sentimen terkait emas juga telah dibuktikan efektivitasnya oleh Norlaila et al., (2024) yang menyelidiki sentimen masyarakat terhadap pertambangan emas di Indonesia melalui Twitter

menggunakan metode K-Means dan Naïve Bayes Classifier. Integrasi kedua algoritma tersebut menghasilkan akurasi gabungan sebesar 99%, yang menunjukkan keefektifannya dalam mengekstraksi dan menganalisis sentimen publik dari data media sosial (Norlaila et al., 2024). Temuan ini semakin memperkuat keyakinan bahwa metode Naïve Bayes merupakan pilihan yang tepat dan teruji dalam analisis sentimen berbasis teks pada topik-topik yang berkaitan dengan emas.

Berdasarkan uraian latar belakang di atas, analisis sentimen masyarakat terhadap investasi emas digital merupakan kajian yang relevan dan penting untuk dilakukan, mengingat pesatnya pertumbuhan investasi digital di Indonesia dan besarnya volume opini masyarakat yang tersebar di media sosial. Algoritma Naïve Bayes telah terbukti efektif dan handal dalam berbagai penelitian terdahulu yang berkaitan dengan topik ini (Ginting et al., 2025; Maulana et al., 2024; Norlaila et al., 2024; Annisa et al., 2024). Oleh karena itu, penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap investasi emas digital menggunakan metode Naïve Bayes, guna memberikan gambaran yang lebih mendalam dan komprehensif mengenai persepsi publik terhadap produk investasi tersebut.

Alasan digunakannya Naïve Bayes dalam penelitian ini tidak lepas dari karakteristik data yang akan diolah, yaitu data teks dari media sosial yang jumlahnya besar dan cenderung tidak terstruktur. Naïve Bayes dikenal memiliki proses komputasi yang ringan dan waktu pelatihan yang singkat, sehingga lebih praktis digunakan untuk mengolah data tweet dalam jumlah banyak dibandingkan metode-metode yang membutuhkan sumber daya komputasi lebih besar. Meskipun algoritma ini didasarkan pada asumsi independensi antar kata yang sebenarnya jarang terpenuhi pada data teks sesungguhnya, pada kenyataannya asumsi tersebut tidak terlalu berpengaruh terhadap hasil klasifikasi dan tetap mampu memberikan akurasi yang cukup baik, bahkan ketika data mengandung banyak singkatan, typo, atau bahasa gaul seperti yang biasa ditemukan di Twitter.

Selain itu, Naïve Bayes juga tidak memerlukan data latih dalam jumlah sangat besar untuk bisa bekerja dengan baik, berbeda dengan pendekatan deep learning yang umumnya baru menunjukkan performa optimal ketika dilatih dengan ribuan hingga jutaan data. Mengingat penelitian ini dilakukan dalam skala skripsi dengan

keterbatasan waktu dan sumber daya, penggunaan Naïve Bayes dirasa lebih realistis untuk diterapkan tanpa harus mengorbankan kualitas hasil analisis. Hal ini juga didukung oleh penelitian-penelitian sebelumnya yang telah membuktikan efektivitas Naïve Bayes pada topik serupa. (Ginting et al., 2025) misalnya, berhasil mengklasifikasikan sentimen masyarakat terhadap emas digital dan memperoleh hasil yang cukup representatif, sementara (Norlaila et al., 2024) bahkan mencatat akurasi hingga 99% ketika Naïve Bayes dikombinasikan dengan algoritma K-Means dalam menganalisis sentimen terkait pertambangan emas. Dengan pertimbangan-pertimbangan tersebut, penelitian ini memilih Naïve Bayes sebagai metode utama dalam melakukan analisis sentimen masyarakat terhadap investasi emas digital.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah:

1. Bagaimana merancang dan mengimplementasikan sistem analisis sentimen berbasis web untuk mengklasifikasikan opini masyarakat terhadap investasi emas digital menggunakan metode Naïve Bayes?
2. Bagaimana tahapan *preprocessing* yang diterapkan pada data komentar media sosial (X/Twitter, Instagram, dan YouTube) sebelum proses klasifikasi sentimen dilakukan?
3. Berapa besar tingkat akurasi, precision, recall, dan F1-score yang dihasilkan oleh metode Naïve Bayes dalam mengklasifikasikan sentimen masyarakat terhadap investasi emas digital ke dalam kategori positif, negatif, dan netral?
4. Bagaimana pengaruh perbandingan rasio pembagian data latih dan data uji (80:20 dan 70:30) terhadap tingkat akurasi model Naïve Bayes yang dibangun?
5. Bagaimana tingkat penerimaan dan kepuasan pengguna terhadap sistem analisis sentimen investasi emas digital yang telah dibangun, berdasarkan hasil pengujian kuesioner?
6. Bagaimana tingkat validitas dan kelayakan hasil pelabelan serta klasifikasi sentimen pada sistem menurut penilaian pakar di bidang *Natural Language Processing* (NLP)?

1.3 Batasan Masalah

Agar penelitian lebih terarah dan tidak menyimpang dari tujuan, maka batasan masalah dalam penelitian ini adalah:

1. Data yang digunakan dalam penelitian ini berupa komentar atau ulasan pengguna terkait investasi emas digital yang diambil dari media sosial, yaitu Instagram mendapatkan data sebanyak 638 data, YouTube mendapatkan data sebanyak 1083 data, dan X (Twitter) mendapatkan 254 data.
2. Data yang dianalisis terbatas hanya pada teks berbahasa Indonesia, sehingga tidak mencakup ulasan dalam bahasa lain.
3. Metode yang digunakan dalam proses klasifikasi adalah algoritma Naïve Bayes.
4. Pembagian data dilakukan dengan rasio 80% Train / 20% Validasi dan 70% Train / 30% Validasi
5. Evaluasi model dilakukan menggunakan confusion matrix, accuracy, precision, recall, dan F1-score.
6. Sistem yang dikembangkan berbasis web dan hanya berfokus pada proses analisis sentimen, tidak mencakup fitur lain di luar analisis.
7. Sistem menyediakan fitur visualisasi kata kunci (keyword) yang menampilkan kata-kata yang paling sering muncul dalam dataset, sehingga dapat diketahui topik atau isu dominan yang dibicarakan masyarakat terkait investasi emas digital.

1.4 Tujuan

Sejalan dengan rumusan masalah yang telah diuraikan, tujuan yang ingin dicapai melalui penelitian ini adalah sebagai berikut:

1. Untuk mengetahui kecenderungan sentimen masyarakat terhadap investasi emas digital.
2. Untuk menerapkan algoritma Naive Bayes dalam proses klasifikasi sentimen teks.
3. Untuk mengukur tingkat akurasi metode Naive Bayes dalam menganalisis sentimen masyarakat.

1.5 Manfaat

Penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. **Manfaat Akademis:** Penelitian ini diharapkan dapat memperkaya pengetahuan dan wawasan di bidang analisis sentimen serta pemrosesan bahasa alami (Natural Language Processing), khususnya dalam konteks analisis opini investor terhadap investasi emas digital yang tersebar di media digital.
2. **Menambah referensi dan literatur ilmiah** di bidang text mining, terutama terkait penerapan algoritma Naive Bayes dalam melakukan klasifikasi sentimen berbasis teks. Penelitian ini juga diharapkan dapat menjadi sumber referensi bagi mahasiswa atau peneliti selanjutnya yang ingin mengembangkan penelitian serupa dengan metode atau objek yang berbeda.
3. **Manfaat Praktis:** Hasil penelitian ini dapat memberikan informasi yang bermanfaat bagi perusahaan atau platform penyedia layanan investasi emas digital sebagai bahan evaluasi terhadap persepsi dan tingkat kepuasan pengguna.
4. **Manfaat bagi Investor dan Masyarakat:** Penelitian ini dapat memberikan gambaran umum mengenai kecenderungan opini masyarakat terhadap investasi emas digital, sehingga dapat membantu calon investor dalam mengambil keputusan investasi secara lebih bijak.

1.6 Sistematika Penulisan

Agar mempermudah pemahaman pada pembahasan penulisan skripsi ini, maka sistematika penulisan diperoleh sebagai berikut:

- BAB I** : Pendahuluan, bab ini membahas latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, serta manfaat penelitian. Bab ini memberikan gambaran umum mengenai perkembangan investasi emas digital di Indonesia melalui platform seperti Pegadaian dan Pluang, serta dilakukannya analisis sentimen masyarakat terhadap layanan tersebut berdasarkan opini yang diperoleh dari media sosial seperti Instagram, YouTube, dan X (Twitter) menggunakan metode Naïve Bayes.
- BAB II** : Tinjauan Pustaka, bab ini berisi tinjauan pustaka yang menjadi landasan teoritis dalam penelitian. Pembahasan meliputi

penelitian-penelitian terdahulu yang relevan terkait analisis sentimen berbasis media sosial, konsep investasi emas digital, analisis sentimen, Natural Language Processing (NLP), preprocessing teks, algoritma Naïve Bayes, pembobotan TF-IDF, serta platform media sosial yang digunakan sebagai sumber data yaitu X (Twitter), Instagram, dan YouTube. Selain itu, bab ini juga membahas teknologi pendukung sistem yang dikembangkan, meliputi bahasa pemrograman Python, database MySQL, web framework Flask, dan evaluasi klasifikasi menggunakan confusion matrix dengan metrik akurasi, presisi, recall, dan F1-score. Uraian pada bab ini bertujuan untuk memberikan dasar pemahaman teoritis dan metodologis yang mendukung proses penelitian serta menjadi acuan dalam pengembangan sistem analisis sentimen yang dibangun.

BAB III : Analisis Dan Perancangan, dalam bab ini menjelaskan analisis dan perancangan sistem yang berbasis web, mencakup kebutuhan fungsional dan non fungsional sistem, perancangan use case diagram, struktur menu, serta flowchart sistem. Selain itu, bab ini menguraikan tahapan pengumpulan data melalui proses crawling dari tiga platform media sosial yaitu X (Twitter), Instagram, dan YouTube, dilanjutkan dengan tahapan preprocessing data teks yang meliputi konversi emoji, cleaning, case folding, normalisasi, tokenizing, stopword removal, dan stemming. Bab ini juga menjelaskan penerapan pembobotan TF-IDF dan algoritma Naïve Bayes dalam proses klasifikasi sentimen ke dalam kategori positif, negatif, dan netral, serta metode evaluasi kinerja model menggunakan confusion matrix dengan metrik akurasi, presisi, recall, dan F1-score. Di bagian akhir, bab ini menyajikan prototype desain antarmuka sistem yang meliputi halaman login, dashboard, data training, data testing, hasil klasifikasi, dan evaluasi model.

BAB IV : Implementasi Dan Pengujian, bab ini membahas hasil implementasi dan pengujian sistem analisis sentimen berbasis web yang telah dikembangkan berdasarkan rancangan pada bab sebelumnya. Pembahasan mencakup tampilan antarmuka sistem mulai dari halaman user, login, dashboard admin, upload dataset, hasil pre-processing, klasifikasi, data testing, hingga halaman cek kalimat. Selain itu, bab ini menyajikan informasi dataset penelitian sebanyak 1978 data yang bersumber dari media sosial X (Twitter), Instagram, dan YouTube dengan pembagian data latih dan data uji. Hasil pengujian metode Naïve Bayes disajikan melalui confusion matrix beserta metrik evaluasi berupa precision, recall, dan F1-score untuk masing-masing kelas sentimen positif, negatif, dan netral.

BAB V Bab ini berisi kesimpulan dari hasil penelitian yang telah dilakukan berdasarkan rumusan masalah dan tujuan penelitian. Selain itu, bab ini juga memuat saran-saran yang dapat diberikan untuk pengembangan penelitian selanjutnya maupun bagi pihak penyedia layanan investasi emas digital dalam meningkatkan kualitas layanan berdasarkan hasil analisis sentimen yang diperoleh.