

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi di era industri 4.0 telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk ekonomi dan keuangan. Salah satu bidang yang terdampak adalah investasi, di mana teknologi memudahkan akses informasi, analisis pasar, dan pengambilan keputusan finansial. Meskipun teknologi memberikan kemudahan, kesadaran Masyarakat terutama di negara berkembang seperti Indonesia terhadap pentingnya investasi masih rendah. Banyak orang lebih memilih pola konsumtif daripada mengalokasikan dana untuk instrumen investasi jangka panjang. Hal ini menjadi tantangan tersendiri mengingat Indonesia memiliki potensi besar dari bonus demografi generasi milenial yang seharusnya dapat menjadi penggerak pertumbuhan ekonomi melalui investasi.

Peningkatan minat investasi masyarakat Indonesia ini dikarenakan kesadaran akan pentingnya menabung dana darurat [1]. Pesatnya digitalisasi, media sosial seperti Twitter menjadi salah satu platform yang banyak digunakan untuk berdiskusi, berbagi opini, dan memperoleh informasi terkait investasi. Data dari Hootsuite (We Are Social, 2024) menunjukkan bahwa Indonesia menempati peringkat kelima sebagai pengguna Twitter terbanyak di dunia. Aktivitas pengguna di platform ini menghasilkan data yang sangat besar, termasuk tweet-tweet yang membahas topik investasi. Data tersebut seringkali tidak dimanfaatkan dengan baik dan hanya menjadi sampah data jika tidak dianalisis lebih lanjut.

Analisis sentimen merupakan cara mengumpulkan pendapat khalayak umum menggunakan jejaring sosial [2]. Analisis sentimen menjadi hal penting. Analisis sentimen adalah cabang dari Natural Language Processing (NLP) yang bertujuan untuk mengidentifikasi dan mengklasifikasikan emosi, pendapat, atau sikap yang terkandung dalam suatu teks, seperti tweet, ulasan, atau komentar. Dengan menggunakan teknik ini, kita dapat memahami persepsi publik terhadap suatu topik dalam hal ini, investasi online secara lebih terstruktur. Analisis sentimen

biasanya mengkategorikan data ke dalam tiga kelas utama: positif, negatif, dan netral. Informasi ini sangat berharga bagi berbagai pemangku kepentingan, seperti pemerintah, pelaku bisnis, dan investor, untuk mengambil kebijakan atau strategi berdasarkan tren opini masyarakat.

Salah satu metode yang efektif untuk analisis sentimen adalah Naïve Bayes Classifier dengan metode probabilistic [3]. Algoritma ini dipilih karena kemampuannya dalam mengolah data teks dengan cepat dan akurat, meskipun dengan dataset yang besar. Naïve Bayes bekerja dengan memanfaatkan probabilitas bersyarat untuk memprediksi kelas suatu dokumen berdasarkan frekuensi kemunculan kata. Keunggulannya terletak pada kesederhanaan komputasi dan ketahanannya terhadap noise dalam data. Untuk meningkatkan akurasi, diperlukan pelabelan manual (manual annotation) namun tidak menutup kemungkinan apabila menggunakan data yang memiliki kapasitas yang banyak perlu untuk pelabelan otomatis dalam tahap pelatihan model. Proses ini melibatkan pemberian label pada data latih agar mesin dapat belajar mengidentifikasi pola sentimen dengan lebih baik.

Penelitian terkait analisis sentiment dalam investasi pernah dilakukan, penelitian tersebut membahas analisis Sentimen Masyarakat Twitter terhadap Emas Digital [4] membahas tentang cryptocurrency [5], tidak hanya itu saja kesenjangan fomo atau fenomena yang terjadi dikalangan Masyarakat berdasarkan isu fomo atau hanya ikut-ikutan tren juga menjadi isu penting dalam analisa opini masyarakat [6], tidak hanya itu juga pembahasa lebih meluas seperti fluktuasi naik turunnya harga saham dalam investasi menjadi peran penting [7].

Peran dalam industri text mining analisis sentimen sering disebut sebagai opinion mining [8]. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat di Twitter terkait investasi online Investasi dalam ekonomi dapat berarti investasi yang dilakukan oleh investor local [9] dengan tujuan menggunakan Naïve Bayes, Hasil analisis diharapkan dapat memberikan gambaran tentang bagaimana publik memandang investasi di era digital apakah cenderung optimis (positif) atau pesimis (negatif). Temuan ini dapat menjadi bahan pertimbangan bagi regulator,

perusahaan fintech, dan edukator keuangan dalam merancang strategi untuk meningkatkan literasi investasi dan minat masyarakat terhadap instrumen keuangan yang produktif.

1.2 Rumusan Masalah

Berdasar latar belakang masalah peneliti menyimpulkan bahwa permasalahan tersebut sebagai berikut:

1. Bagaimana algoritma Naïve bayes multinomial dalam melakukan klasifikasi pada data twitter tentang investasi ?
2. Berapa tingkat performa algoritma naïve bayes multinomial yang dihasilkan ?
3. Bagaimana Algoritma Naïve bayes agar dapat bekerja untuk melakukan analisis sentiment tentang investasi?

1.3 Batasan Masalah

Tuliskan Batasan Masalah yang ada pada penelitian, peneliti memberi suatu Batasan terkait penelitian ini, diwaktu melakukan proses penelitian berlangsung sebagai berikut:

1. tweet akan diambil dengan kata kunci yang relevan dengan topik penelitian yaitu tentang investasi di Indonesia yaitu saham dan investasi.
2. Tweet diambil dengan pada Januari tahun 2024 hingga Desember 2024.
3. Algoritma yang digunakan hanya akan menggunakan algoritma naïve bayes multinomial.
4. Data Peneliti tidak melakukan perbandingan dengan algoritma lain dengan topik klasifikasi yang menggunakan metode berbeda.
5. Peneliti hanya akan mengambil data pada sosial media twitter dan tidak melakukan penambahan data pada media sosial lain kecuali twitter.
6. Tweet akan diubah ke dalam satu bahasa yaitu bahasa Indonesia.

1.4 Tujuan Penelitian

Tujuan Peneliti dari penelitian ini merupakan tujuan bentuk riset dengan sebagai berikut:

1. Membangun machine learning dan mengolah data media sosial menggunakan metode naïve bayes.
2. Peneliti melakukan uji coba pada metode naïve dalam mengetahui pendapat masyarakat dalam berinvestasi.
3. Melihat dampak investasi melalui platform twitter.
4. Mendapat ulasan penting tentang investasi.

1.5 Sistematika Penulisan

Sistematika penulisan skripsi yang akan digunakan dalam penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN, Berisi urgensi penelitian mengenai opini publik terhadap investasi digital yang semakin berkembang di era teknologi. Latar belakang diuraikan dengan mengangkat fenomena maraknya diskusi publik di media sosial mengenai investasi digital, serta perlunya analisis otomatis menggunakan pendekatan machine learning. Rumusan masalah, tujuan, manfaat, dan batasan penelitian dijabarkan untuk memperjelas arah kajian.

BAB II TINJAUAN PUSTAKA, Berisi teori dan konsep terkait opini publik, investasi digital, serta pengolahan bahasa alami (Natural Language Processing/NLP). Selain itu, dibahas pula teori mengenai algoritma Multinomial Naïve Bayes sebagai metode klasifikasi teks, serta tinjauan terhadap penelitian sejenis yang mendukung landasan ilmiah.

BAB III METODE PENELITIAN, Berisi penjabaran tahapan teknis dalam penelitian, mulai dari pengumpulan data opini publik (misalnya dari Twitter, forum online, atau ulasan), preprocessing teks (tokenisasi, stopword removal, stemming), pembobotan (TF-IDF atau lainnya), hingga penerapan algoritma Multinomial Naïve Bayes untuk klasifikasi opini

(positif, negatif, netral). Disertakan pula evaluasi performa model seperti akurasi, presisi, recall, dan f1-score.

BAB IV HASIL DAN PEMBAHASAN, Menyajikan hasil pengolahan dan klasifikasi opini publik terhadap investasi digital. Disertai analisis terhadap distribusi sentimen, efektivitas model Naive Bayes dalam memahami konteks opini, serta interpretasi data terhadap kondisi sosial yang terkait. Pembahasan juga mengaitkan hasil dengan studi terdahulu dan konteks kebijakan atau tren digital saat ini.

BAB V PENUTUP, Memuat simpulan dari temuan utama penelitian serta saran untuk pengembangan metode atau perluasan data di masa mendatang. Penutup juga mencerminkan kontribusi penelitian terhadap bidang teknologi informasi dan pengambilan keputusan berbasis opini publik.