

OAJIS

Open Access
Journal of
Information
Systems

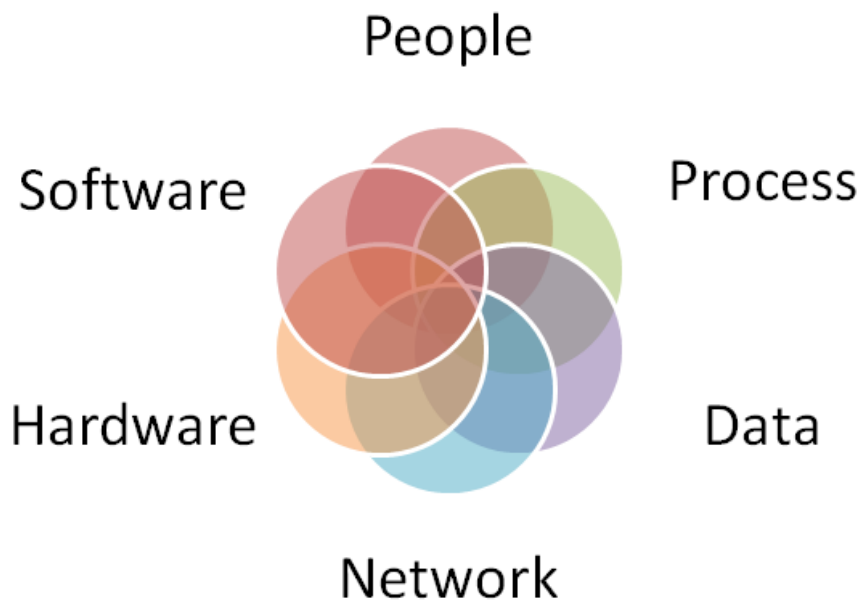
is.its.ac.id/pubs/oajis/

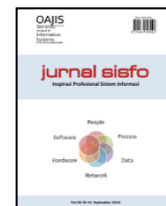
ISSN 1979-3979



jurnal sisfo

Inspirasi Profesional Sistem Informasi





Pimpinan Redaksi

Sholiq (Institut Teknologi Sepuluh Nopember)

Dewan Redaksi

Reny Nadlifatin (Institut Teknologi Sepuluh Nopember)

Tining Haryanti (Universitas Muhammadiyah Surabaya)

Faizal Mahananto (Institut Teknologi Sepuluh Nopember)

Rizal Risnanda Utama (Institut Teknologi Sepuluh Nopember)

Dimas Agung Perkasa (Institut Teknologi Sepuluh Nopember)

Monica Widiarsi (Universitas Surabaya)

Anjik Sukmaaji (Universitas Dinamika)

Devi Septiani (Universitas Brawijaya)

Tata Pelaksana Usaha

Rachmatina Retno Septiani

Sekretariat

Departemen Sistem Informasi – Fakultas Teknologi Elektro dan Informatika Cerdas

Institut Teknologi Sepuluh Nopember (ITS) – Surabaya

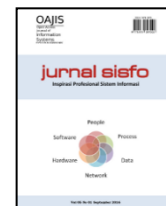
Telp. 031-5999944 Fax. 031-5964965

Email: editor@jurnalsisfo.org

Website: <http://jurnalsisfo.org>

Jurnal SISFO juga dipublikasikan di *Open Access Journal of Information Systems* (OAJIS)

Website: <http://is.its.ac.id/pubs/oajis/index.php>



Mitra Bestari

Prof. Nur Aini Rakhmawati, S.Kom., M.Sc.Eng., Ph.D
(Institut Teknologi Sepuluh Nopember)

Dr. Jusak (James Cook University, Singapore)

Dr. Muhammad Ainul Yaqin, S.Si., M.Kom. (UIN Maulana Malik Ibrahim)

Dr. Bambang Setiawan, S.Kom., M.T. (Institut Teknologi Sepuluh Nopember)

Dr. Feby Artwodini, S.Kom., M.T. (Institut Teknologi Sepuluh Nopember)

Arif Wibisono, S.Kom., M.Sc., Ph.D. (Institut Teknologi Sepuluh Nopember)

Dr. Agus Subhan Akbar, S.Kom, M.Kom. (Universitas Islam Nahdlatul Ulama Jepara)

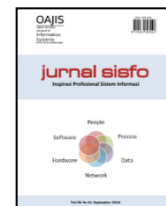
Ir. Nunik Endah Sulistiyawati, M.T. (TNI AL, Indonesia)

Dhiani Tresna Absari, S.T. M.Kom. (Universitas Surabaya)

Muhamad Amirul Haq, S.T., M.Sc. (Universitas Muhammadiyah Surabaya)

Ronny Trian Surbakti S.IP., M.M. (Universitas Katolik Parahyangan)

Agus Dwi Purwolastono, SE, M.Acc, Ak. (Institut Teknologi Sepuluh Nopember)



Daftar Isi

Analisis Sentimen dan Korelasi Berita Saham pada Platform Trading Gotrade terhadap Perubahan Harga Saham <i>Christoforus Fachrasya, Retno Aulia Vinarti, Faizal Mahananto, Amalia Utamima, Renny Pradina Kusumawardani</i>	1
Proyek Pengembangan Notares Website Manajemen Keuangan Cerdas dengan Metode Scrum Agile <i>Rahmat Ramadhan Permana, Agung Budi Prasetya, Farhan Adika Suwardana, Arayzi Rayyansyah, Arjuna Putra Kharisma, Sholiq Sholiq, Rizal Risnanda Hutama</i>	12
A Bibliometric Analysis of Digital Business Models: Comparative Insights Before and After COVID-19 <i>Febby Candra Pratama</i>	32
Mapping a Decade of Digital Twin Research: Trends, Thematic Evolution, and China's Strategic Lead <i>Radityo Prasetyanto Wibowo, Prasasti Karunia Farista Ananto, Eugenia Indrawan</i>	46
Pengaruh Fitur Social Media Marketing terhadap Keputusan Pembelian Brand Fashion <i>Mudjahidin Mudjahidin, Ahmad Ikhsan, Talitha Firyal Ghina Nuha</i>	61
Penyusunan Disaster Recovery Planning (DRP) Pada Data Center Unusa Menggunakan NIST 800-34 <i>Riko Adi Kurniawan, Endang Sulistiyani</i>	81
Application of Gamification to Enterprise Systems: A Systematic Literature Review <i>Tri Puspa Rinjeni, Mahendrawathi ER</i>	97



Halaman ini sengaja dikosongkan

Analisis Sentimen dan Korelasi Berita Saham pada *Platform Trading* Gotrade terhadap Perubahan Harga Saham

Christoforus Fachrasya, Retno Aulia Vinarti*, Faizal Mahananto, Amalia Utamima,
Renny Pradina Kusumawardani

Departemen Sistem Informasi, Fakultas Teknologi Elektro dan Informatika Cerdas, Institut Teknologi Sepuluh Nopember

Abstract

The interest in investing in stocks has increased in recent years. Gotrade app launched to address this increasing interest. Gotrade Daily is a newsletter that provides recent information about stocks that are traded in Gotrade application. Effectiveness of the Gotrade Daily newsletter as the main reference for investors is the focus on this research. The effectiveness of Gotrade Daily newsletter can be calculated by measuring the correlation between its news sentiment and stock price changes. News sentiment is calculated using rule-based approach by combining two libraries, Loughran-McDonald dictionary and VADER (Valence Aware Dictionary and sEntiment Reasoner) dictionary. The obtained sentiment scores could be correlated with its stock price changes, using Pearson correlation and Spearman correlation. From the results of several experiments that have been conducted, it was concluded that there is almost no connection between stock news and the closing value of the following day. So that the news published in the Gotrade Daily newsletter cannot be relied on as the only consideration for making decisions to invest in certain stocks.

Keywords: Sentiment analysis, Stock price, Loughran-McDonald, VADER, Pearson correlation, Spearman correlation

Abstrak

Investasi saham dalam beberapa tahun terakhir mengalami peningkatan. Aplikasi Gotrade menjawab peningkatan minat investasi tersebut. Gotrade Daily, merupakan newsletter berisi informasi-informasi terkini mengenai saham-saham yang ada di aplikasi Gotrade. Efektivitas *newsletter* Gotrade Daily sebagai referensi bagi investor inilah yang menjadi bahasan pada penelitian ini. Salah satu cara untuk mengukur efektivitas adalah dengan cara menghitung korelasi antara sentimen berita terhadap perubahan harga saham. Korelasi diukur dengan membandingkan nilai sentimen berita dengan perubahan harga saham. Analisis sentimen pada data berita menggunakan pendekatan rule-based, dengan memanfaatkan kombinasi library Loughran-McDonald dictionary dan VADER (Valence Aware Dictionary and sEntiment Reasoner) *dictionary*. Data berita yang sudah didapatkan nilai analisis sentimen kemudian dikorelasikan dengan perubahan harga saham. Pengkorelasiannya ini didapatkan dengan memanfaatkan korelasi Pearson dan korelasi Spearman. Dari hasil beberapa percobaan yang telah dilakukan didapatkan kesimpulan bahwa antara berita saham dan nilai closing hari setelahnya hampir tidak memiliki keterhubungan. Sehingga berita yang dimuat pada *newsletter* Gotrade Daily tidak dapat diandalkan sebagai satu-satunya pertimbangan untuk mengambil keputusan untuk berinvestasi ke saham tertentu.

Kata kunci: Analisis Sentimen, Harga Saham, Loughran-McDonald, VADER, Korelasi Pearson, Korelasi Spearman

*Corresponding Author

Email address: ra_vinarti@its.ac.id (Retno Aulia Vinarti)

<https://doi.org/10.24089/j.sisfo.2025.07.001>

1. Pendahuluan

Investasi saham menjadi hal yang marak dalam beberapa tahun terakhir. Melalui *press release*, Bursa Efek Indonesia (BEI) menyatakan bahwa ada 6 juta *single investor identification* (SID) pada September 2024, tepatnya adalah sebesar 6.001.573. BEI juga memaparkan bahwa pada tahun 2024 ada 744 ribu investor saham baru, mencatatkan kurang lebih 12% pertumbuhan [1].

Saham adalah tanda penyertaan modal suatu pihak dalam suatu perusahaan. Saham juga merupakan instrumen pasar keuangan yang digunakan perusahaan untuk mendanai operasionalnya. Saham bagi investor, merupakan instrumen investasi yang banyak dipilih karena dividen (keuntungan) dan capital gain yang menarik. Dibalik keuntungannya, saham memiliki risiko, seperti *capital loss* dan risiko likuidasi atau kebangkrutan yang membawa kerugian bagi investor [2]. Dalam melakukan transaksi saham, harga saham merupakan hal penting yang harus diperhatikan investor. Harga saham yang cenderung naik akan meningkatkan minat investor berinvestasi pada saham tersebut, harga saham yang cenderung turun akan menurunkan minat investor untuk berinvestasi pada saham tersebut. Harga saham sendiri dipengaruhi oleh beberapa hal, salah satunya adalah berita. Berita merupakan salah satu sumber yang digunakan oleh para investor untuk membantu mengambil keputusan untuk membeli saham atau tidak. Berita dengan sentimen yang baik mendorong minat investor untuk berinvestasi, sedangkan berita yang buruk menurunkan minat investor untuk berinvestasi [3].

Gotrade adalah platform trading yang menawarkan investor untuk berinvestasi di pasar saham Amerika Serikat. Gotrade didirikan pada tahun 2019 sebagai solusi bagi investor dari mana saja di seluruh dunia untuk berinvestasi di pasar saham di Amerika Serikat, dalam hal ini bursa saham NYSE (New York Stock Exchange) dan NASDAQ (National Association of Securities Dealers Automated Quotations). Sesuai dengan misi dari Gotrade yaitu, “*Make investing accessible for literally anyone*”, Gotrade mengupayakan berbagai cara untuk memudahkan investor dalam berinvestasi, seperti memungkinkan investor untuk berinvestasi saham hanya mulai dari \$1 (Rp. 15.000), dan menyediakan *newsletter* bernama *Gotrade Daily* sebagai referensi bagi investor dalam pengambilan keputusan dalam berinvestasi pada salah satu saham. Hal inilah yang menjadi pertanyaan apakah berita perusahaan yang dimuat pada *Gotrade Daily* berkorelasi dengan naik turunnya harga saham perusahaan tersebut.

Wang et al. [4], melalui studi kasusnya, menemukan adanya korelasi antara sentimen berita terhadap berubahnya harga saham. Untuk menganalisis sentimen berita, dapat menggunakan analisa yang dikenal sebagai analisis sentimen. Analisis sentimen (disebut juga *opinion mining*) adalah ilmu yang menganalisis pendapat, sentimen, evaluasi, penilaian, sikap, dan emosi seseorang terhadap suatu entitas seperti, produk, layanan, organisasi, individu, isu, topik, dan atributnya [5]. Analisis sentimen merupakan hal yang kompleks, melibatkan 5 tahapan yang berbeda untuk menganalisis data sentimen yaitu *data collection*, *text preparation*, *sentiment detection*, *sentiment classification*, dan *presentation of output*. Ada beberapa pendekatan untuk melakukan analisis sentimen, yaitu *machine learning*, *rule-based*, dan *hybrid*. Analisis sentimen digunakan di berbagai bidang seperti, pemasaran, politik, dan sosiologi [6].

Ada beberapa usulan untuk memberikan gambaran pengaruh berita terhadap berubahnya harga saham. Alamsyah et al. [7], melalui analisis korelasi terhadap 20 perusahaan yang ada pada indeks saham LQ45 dengan sentimen berita, berpendapat bahwa korelasi antara sentimen berita dan perubahan harga saham adalah kurang kuat. Az-Zahra et al. [8], melakukan analisis korelasi pada media-media sosial (Facebook, Twitter, dan Instagram) dan media berita, memberikan kesimpulan bahwa media sosial Facebook dan Twitter, serta berita dapat menjadi acuan investor untuk mengambil keputusan berinvestasi pada saham

perusahaan otomotif. Rai [9] menyimpulkan lewat penelitian korelasi antara pemberitaan vaksin Covid-19 dan perubahan harga saham emiten farmasi KLBF (PT Kalbe Farma Tbk), KAEF (PT Kimia Farma Tbk), INAF (PT Indofarma Tbk), dan PEHA (PT Phapros Tbk), bahwa pada perubahan harga saham KLBF ada hubungan korelasi searah yang signifikan dengan sentimen pemberitaan, sedangkan pada perusahaan lainnya tidak menunjukkan hubungan yang signifikan dengan sentimen pemberitaan.

Melihat dari permasalahan tersebut, maka perlu untuk dilakukan penelitian bagaimana korelasi antara berita yang dibahas dengan perubahan harga saham setelah berita dimuat. Melalui penerapan *natural language processing*, analisis sentimen, serta korelasi *Pearson* dan korelasi *Spearman*, maka penelitian ini dapat menjawab permasalahan tersebut, menunjukkan waktu sentimen berita berimbas terhadap perubahan harga saham, dan kekuatan korelasi berita pada newsletter *Gotrade Daily* dengan perubahan harga saham.

2. Tinjauan Pustaka

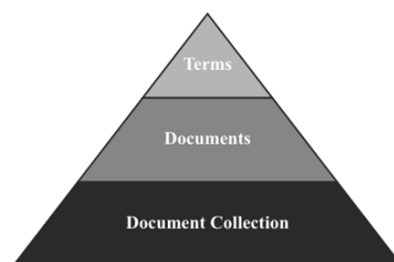
Penelitian ini menggunakan beberapa landasan teori yang relevan dengan tujuan penelitian, yaitu *Natural Language Processing* (NLP), *Text Preprocessing*, *Sentiment Analysis*, *VADER* dan *Loughran-McDonald*, *Spearman and Pearson Correlation*, Saham dan *Platform Trading GoTrade*.

2.1 Pengolahan Bahasa Alami

Pengolahan bahasa alami, atau yang biasa disebut dengan *Natural Language Processing* (NLP) adalah bidang ilmu yang berfokus pada memahami interaksi komputer dengan bahasa alami manusia. Bahasa alami (*Natural Language*) yang dimaksud disini adalah bahasa yang biasa dituturkan oleh manusia. Bidang ilmu ini memiliki tujuan agar komputer mampu memahami perintah yang diberikan oleh manusia dalam bahasa alaminya, dan bertindak sesuai dengan perintah tersebut [10]. Pengolahan bahasa alami memiliki beberapa tingkatan pemahaman, yang mana dimulai dari tingkat yang paling luas hingga paling sempit, yaitu *phonology*, *morphology*, *lexical*, *syntactic*, *semantic*, *discourse*, dan *pragmatic* [11].

2.2 Praproses Teks

Praproses teks adalah teknik untuk menyiapkan data teks sehingga mampu digunakan untuk analisis. Praproses teks menguraikan kumpulan data teks sehingga didapatkan identitas atau *term* pada setiap kumpulan data teks tersebut [12]. Gambar 1 mengilustrasikan hierarki dari komponen pada data teks.



Gambar 1 Hierarki Komponen Teks

Praproses memiliki beberapa tahapan yang meliputi *Noise Cleaning* yaitu membersihkan teks dari kata yang tidak relevan (*noise*), *Case Folding* yang mengubah bentuk kata menjadi kecil atau *lowercase*, kemudian teks dipecah menjadi bagian-bagian kecil melalui *Tokenizing*, kata yang tidak signifikan dihapus dengan *Stopwords Removal*, dan *Stemming* atau *Lemmatization* yang mengubah kata menjadi bentuk dasarnya. Teknik praproses dapat menggunakan *python library* seperti *spaCy* atau *NLTK* [13].

2.3 Analisis Sentimen

Analisis sentimen merupakan penggunaan *text analysis*, dan *computational linguistics* untuk memahami, mengekstrak dan mengolah data tekstual secara otomatis sehingga informasi sentimen yang terkandung dalam suatu opini bisa didapatkan. Analisis sentimen bertujuan untuk menganalisis polaritas sentimen atau opini dari pelanggan terhadap suatu entitas atau objek. Analisis sentimen memiliki banyak nama. Sering kali analisis sentimen disebut sebagai *subjectivity analysis*, *opinion mining*, atau *appraisal extraction* [5]. Analisis sentimen memiliki beberapa pendekatan untuk menyelesaikan permasalahannya. Pendekatan itu antara lain adalah pendekatan *rule-based*, pendekatan *Machine Learning*, dan pendekatan *Hybrid*. *Rule-based*, yang disebut juga sebagai *dictionary-based* atau *lexicon-based*, merupakan pendekatan yang mengklasifikasikan kata-kata dalam sebuah “kamus” polaritas dari kata tersebut (positif atau negatif). Cara kerjanya adalah setiap kata mulanya akan bernilai nol, dan akan diberikan nilai. Nilai dari setiap kata bervariasi, tergantung setiap peneliti menerapkan pendekatan ini. Contohnya nilai dari kata diklasifikasikan dalam 5 kategori, yaitu sangat positif, positif, netral, negatif, dan sangat negatif. Ada juga yang memberikan nilai menggunakan skala 10 poin, dimulai -5 hingga +5. Pendekatan *machine learning* merupakan pendekatan yang menggunakan algoritma untuk membangun sebuah *classifier* sebagai penentu polaritas dari sebuah teks. *Classifier* ini dibuat dengan melakukan *training* pada data yang sudah dilabeli sentimennya. Banyak algoritma *machine learning* yang bisa digunakan, namun yang algoritma yang populer digunakan adalah *Support Vector Machine (SVM)*, *Naive Bayes*, *N-gram*, *K-Nearest Neighbour*, dan *Neural Network*. Pendekatan *Hybrid* menggabungkan beberapa metode analisis sentimen, baik dari *Machine Learning* maupun *Rule-Based*. Pendekatan ini biasanya digunakan untuk menggabungkan kelebihan metode analisis sentimen, dan juga mengatasi limitasi yang muncul pada metode analisis sentimen [14].

2.4 VADER dan Loughran-McDonald

VADER (*Valence Aware Dictionary and sEntiment Reasoner*) adalah alat analisis sentimen berbasis *rule-based* yang didesain secara khusus untuk teks singkat atau teks sosial media. VADER memiliki beberapa fitur yaitu memiliki leksikon kata yang sudah diberi skor sentimen berkisar antara -4 (sangat negatif) hingga +4 (sangat positif), mampu mengenali ragam bahasa informal, slang, hingga emotikon. VADER juga mampu membedakan kata, puntuasi, hingga kapitalisasi yang memperkuat sentimen pada sebuah teks [15]. *Loughran-McDonald dictionary* adalah alat analisis sentimen berbasis *rule-based* yang berspesialisasi pada teks finansial. *Loughran-McDonald dictionary* dikembangkan untuk mengatasi kesalahan deteksi sentimen pada teks finansial yang sering terjadi pada alat analisis sentimen konvensional. Leksikon kata pada alat ini berbasis pada laporan 10K tahunan milik perusahaan publik yang wajib diserahkan kepada Komisi Sekuritas dan Bursa (SEC) Amerika Serikat. Spesialisasi *Loughran-McDonald dictionary* ini yang membuat analisis sentimen lebih mampu memahami konteks yang lebih akurat dan reliabel pada teks finansial seperti berita finansial atau laporan keuangan [16]. Kelebihan dari kedua analisis sentimen ini menjadi dasar sebuah penelitian untuk bereksperimen dengan cara mengombinasikan kedua analisis sentimen ini pada berita finansial. Penelitian ini menyatukan VADER dan *Loughran-McDonald* dengan cara menggabungkan kedua leksikon. Penelitian ini menghasilkan akurasi sebesar 59,7% dan menyimpulkan bahwa gabungan kedua alat ini masih kurang cukup untuk meningkatkan ketepatan analisis sentimen [17].

2.5 Korelasi Pearson dan Spearman

Pearson Correlation adalah sebuah metode yang mengukur seberapa kuatnya hubungan antara dua variabel. *Pearson Correlation* mengukur hubungan kedua variabel dengan menggunakan representasi angka -1, 0, atau 1 pada setiap variabel dengan formula sebagai berikut:

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (1)$$

Dimana, r adalah korelasi *pearson*, n adalah jumlah data x adalah nilai variabel x , y adalah nilai variabel y , dan \sum adalah notasi jumlah. Bila angka korelasi mendekati 1 menandakan hubungan korelasi selaras yang cenderung kuat, dan sebaliknya bila angka korelasi mendekati -1 menandakan hubungan korelasi bertolak belakang yang cenderung kuat, bila angka korelasi mendekati 0 menandakan hubungan korelasi yang lemah atau tidak ada [18].

Spearman Correlation sebuah metode yang mengukur seberapa kuatnya hubungan antara dua variabel berdasarkan peringkat. *Spearman Correlation* menghitung hubungan data secara monotonik, cocok untuk mengukur korelasi antar data yang bersifat ordinal. Korelasi Spearman dihitung dengan cara memberikan peringkat pada data, kemudian diterapkan dengan formula sebagai berikut:

$$\rho = \frac{6 \sum d_i^2}{n(n^2-1)} \quad (2)$$

Dimana, ρ adalah korelasi *spearman*, d_i adalah selisih peringkat antara variabel untuk data ke- i , dan n adalah jumlah data [19].

2.6 Saham dan Platform Trading Gotrade

Saham adalah bukti penyertaan modal atas sebuah perusahaan. Kepemilikan atas saham merupakan bukti investasi dana atas sebuah perusahaan, dan kepemilikan dengan besaran tertentu atas perusahaan tersebut [20]. Memiliki saham memiliki beberapa manfaat, seperti mampu memberikan partisipasi dalam pengambilan keputusan internal perusahaan melalui Rapat Umum Pemegang Saham (RUPS), potensi mendapatkan deviden atau yang disebut juga pembagian hasil laba perusahaan [21].

Gotrade adalah *platform* investasi berbasis di Singapura yang memberikan akses kepada pengguna untuk berinvestasi saham perusahaan-perusahaan besar yang melantai di bursa saham Amerika Serikat (NYSE dan NASDAQ). Kelebihan *Gotrade* adalah memungkinkan pengguna mengakses dan berinvestasi saham-saham tersebut dari manapun pengguna berada, hanya dengan modal sebesar US\$ 1. *Gotrade* didirikan pada tahun 2019 oleh Rohit Mulani, Norman Wanto, dan David Grant. *Gotrade* mendapatkan investasi *seed-round* sebesar US\$ 7 juta pada 2021 dari *LocalGlobe* dan *Social Leverage*, yang juga merupakan investor awal dari *platform trading* yang serupa yaitu *Robinhood*. *Gotrade* diluncurkan secara tertutup pada Maret 2021, yang kemudian diluncurkan secara global pada September 2021. *Gotrade* hingga saat ini sudah memiliki 500,000 pengguna, dengan 90 persen dari para pengguna tersebut berinvestasi di pasar saham Amerika untuk pertama kalinya. Pengguna *Gotrade* hingga saat ini mampu untuk melakukan transaksi yang bisa lebih dari US\$ 400 juta dengan. Saat ini *Gotrade* berhasil mengumpulkan US\$ 15,5 juta dalam putaran seri A yang dipimpin *Velocity Capital Fintech Ventures* yang membuat total pendanaan *Gotrade* saat ini mencapai US\$ 22,5 juta [22]. Penelitian-penelitian terdahulu yang terkait dengan platform GoTrade juga menjadi rujukan dari penelitian ini. *Platform GoTrade* telah diulas secara global dan menghasilkan *review* yang baik pada NASDAQ, London, dan HongKong exchanges. Walaupun demikian terdapat pengecualian-pengecualian seperti beberapa biaya yang terasosiasi saat melakukan transaksi jual-beli dengan platform ini [23].

3. Metodologi

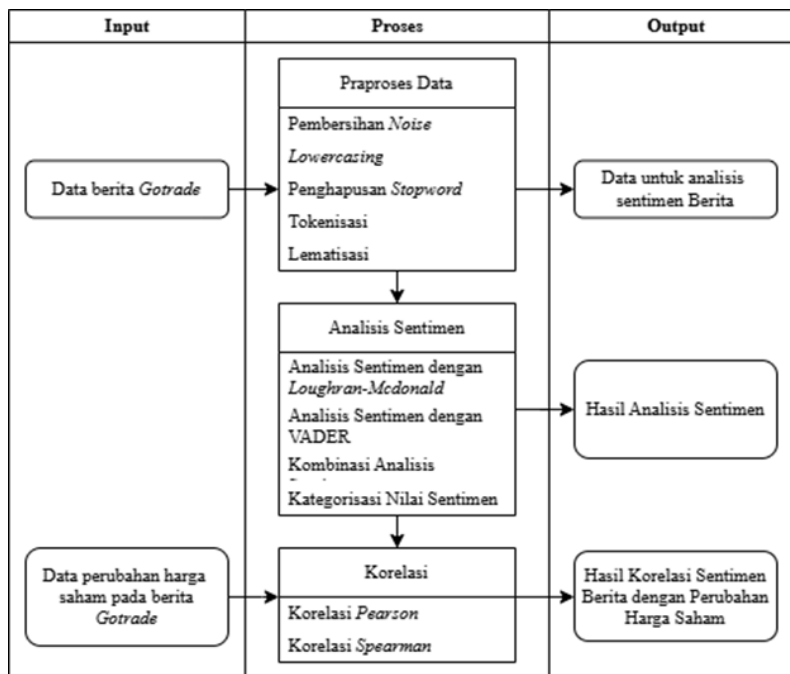
Penelitian dilakukan melalui beberapa tahapan, yaitu analisis permasalahan, studi literatur, pengumpulan data, analisis sentimen, korelasi berita dengan harga saham, uji coba dan pengambilan kesimpulan. Tahapan digambarkan dalam bentuk diagram sebagai berikut.



Gambar 2 Diagram Tahapan Penelitian

Pengumpulan data mengumpulkan berita-berita yang dimuat pada situs *Gotrade Daily* (<https://gotrade.substack.com/>), dari tanggal 11 Mei 2021 hingga 24 Maret 2022, menggunakan ekstensi perambanan *WebScraper*. Data yang dikumpulkan meliputi tanggal (*date*), dan judul berita (*title*). Setelah data berita didapatkan, dikumpulkan data saham yang akan digunakan sebagai bahan penelitian. Data yang dikumpulkan adalah data saham yang dimuat pada berita dari tanggal 11 Mei 2021 hingga 24 Maret 2022, dan juga diperdagangkan di aplikasi *Gotrade*. Data saham dikumpulkan dengan *software Microsoft Excel*, menggunakan fungsi `STOCKHISTORY()`, pada waktu harga saham ditutup pada hari berita dimuat (*dclosing*), harga saham dibuka 1 hari kerja setelah berita dimuat (*d+1 opening*), harga saham ditutup 1 hari kerja setelah berita dimuat (*d+1 closing*), harga saham dibuka 2 hari kerja setelah berita dimuat (*d+2 opening*), dan harga saham ditutup 2 hari kerja setelah berita dimuat (*d+2 closing*).

Data yang telah dikumpulkan kemudian akan masuk ke dalam model analisis sentimen dan korelasi. Model analisis sentimen dan korelasi dirancang menggunakan *jupyter notebook* pada aplikasi *web Google Colab* (<https://colab.research.google.com/>). Perancangan model dibuat dengan tahapan-tahapan pada Gambar 3.



Gambar 3 Diagram Tahapan Analisis Sentimen dan Korelasi

Untuk meningkatkan akurasi dan konsistensi, sebelum analisis sentimen dilakukan praproses teks pada data yang sudah dikumpulkan. Langkah-langkah praproses yang dilakukan meliputi penghapusan *emoji*, *lowercasing*, tokenisasi, penghapusan *stopwords*, dan lematisasi.

Analisis sentimen berita menggunakan kombinasi analisis dua *rule-based* yaitu *Loughran-McDonald* dan *VADER*. Data berita untuk *Loughran-McDonald Dictionary* dimuat ke dalam model, kemudian membagi data menjadi dua himpunan variabel. Kata-kata yang bernilai lebih besar dari 0 pada kolom *positive*, masuk pada variabel *positive_words*, kata-kata yang memiliki nilai lebih besar dari 0 pada kolom *negative*, masuk pada variabel *negative_words*. Analisis Sentimen dengan *VADER* memuat *SentimentIntensityAnalyzer()* dengan membuat variabel *vader_analyzer*. Fungsi ini dipanggil dari *library vaderSentiment* yang telah diinstal sebelumnya, yang kemudian mengimpor *SentimentIntensityAnalyzer* dari *vaderSentiment*. Kedua hasil analisis sentimen kemudian dikombinasikan dengan dengan menjumlahkan hasil pada kolom

tokens_sw untuk analisis sentimen *Loughran-McDonald* dan kolom *lemma* untuk analisis sentimen VADER, dengan pembobotan masing-masing 0,5. Penerapan pada kolom yang berbeda ini dikarenakan *Loughran-McDonald* tidak memerlukan teks berita yang dilematisasi (kolom *lemma*), sedangkan VADER memerlukan teks berita yang dilematisasi.

Hasil analisis sentimen kemudian dikorelasikan dengan perubahan nilai saham dengan menggunakan korelasi *pearson* dan korelasi *spearman*. Skor korelasi *pearson* bila mendekati 1 menandakan hubungan korelasi yang cukup kuat, bila skor mendekati -1 menandakan hubungan bertolak belakang yang cukup kuat, bila skor mendekati 0 menandakan hubungan yang cukup lemah. Korelasi *Pearson* dihitung menggunakan *pearsonr* yang diimpor dari *library scipy.stats*. Pada penghitungan korelasi digunakan fungsi *pearsonr*. Korelasi *Spearman* dihitung menggunakan *spearmanr* yang diimpor dari *library scipy.stats*. Penghitungan korelasi menggunakan fungsi *spearmanr* yang telah diimpor sebelumnya. Nilai hasil korelasi dibulatkan ke angka desimal keempat.

4. Hasil dan Pembahasan

4.1 Hasil Analisis Sentimen

Berita-berita yang dimuat pada *Gotrade Daily* pada rentang waktu 11 Mei 2021 hingga 24 Maret 2022 adalah data berita yang dikumpulkan. Tabel 1 adalah sampel hasil pengumpulan data berita pada *Gotrade Daily*.

Tabel 1. Sampel Hasil Pengumpulan Data Berita

Date	Title
Mar 24, 2022	Ouch! Shares of Adobe are down nearly 10% after reporting earnings 😞
Mar 24, 2022	☑️ Spotify will now be allowed to offer its own billing on the Play Store 📺
Mar 24, 2022	📉 “Buy the dip and stay long on Nvidia” - Jim Cramer ✅
Mar 23, 2022	Ryan Cohen sends GameStop shares rallying 🚀
Mar 23, 2022	☑️ Apple’s services are back up and running after outage 👍
Mar 23, 2022	☑️ South Korea approves Merck’s Covid pills 🍬
Mar 22, 2022	Alibaba shares are up nearly 9% after announcing record share buyback program of \$25 billion 🚀
Mar 22, 2022	☑️ Tesla’s German Gigafactory is finally LIVE! 🇩🇪
Mar 22, 2022	☑️ Nike’s shares are up nearly 6% after reporting earnings 📈
Mar 21, 2022	Berkshire Hathaway plans to buy Alleghany for \$11.6 billion 💰









Data judul berita yang sudah dikumpulkan kemudian dilakukan pembersihan, tokenisasi, penghapusan *stopwords*, dan lematisasi. Contoh *stopwords* yang akan dihapus pada konteks berita saham ini adalah *will*, *are*, dan lain-lain. Penghapusan ini menjadi satu tahapan yaitu praproses data untuk mempersiapkan data-data judul berita sehingga layak dan muda untuk dianalisis. Kolom *tokens_sw* adalah tokens yang telah dihapus *stopwords*-nya. Sedangkan kolom *lemma* berisi tokens yang telah dikembalikan ke kata dasarnya. Sampel data hasil praproses data berita ditampilkan pada Tabel 2.

Data berita yang telah dipraproses dibuat analisis sentimen menggunakan pendekatan *rule-based* dengan mengombinasikan *library VADER* (*Valence Aware Dictionary and sEntiment Reasoner*) dan *library Loughran-McDonald Master Dictionary*. Nilai polaritas sentimen didapatkan dengan cara menjumlahkan nilai sentimen dari *library VADER* dan *library Loughran-McDonald*, yang diberi bobot masing-masing 0,5.



Nilai polaritas sentimen yang sudah didapatkan kemudian dikategorisasi ke dalam tiga kelompok, yaitu positif, netral, dan negatif. Masing-masing kelompok direpresentasikan dalam bentuk angka. Jika nilai polaritas sentimen lebih besar dari 0 ($n > 0$), maka nilai sentimen tersebut termasuk positif, bernilai 1 dan 3. Jika nilai polaritas sentimen sama dengan 0 ($n = 0$), maka nilai sentimen tersebut termasuk netral, bernilai 0 dan 2. Jika nilai polaritas sentimen lebih kecil dari 0 ($n < 0$), maka nilai sentimen tersebut termasuk negatif, bernilai -1 dan 1. Nilai -1, 0, dan 1 masuk ke dalam kolom *sentiment*, sebagai variabel korelasi *pearson*.

Nilai 1, 2, dan 3 masuk ke dalam kolom *rank*, sebagai variabel korelasi *spearman*. Tabel 3 adalah tabel sampel hasil analisis sentimen.

Tabel 2. Sampel Hasil Praproses Data Berita

	title	title_clean	tokens	tokens_sw	lemma
0	Ouch! Shares of Adobe are down nearly 10% after...	ouch shares of adobe are down nearly 10 after ...	[ouch, shares, of, adobe, are, down, nearly, 1...	[ouch, shares, adobe, down, nearly, 10, report...	[ouch, share, adobe, down, nearly, 10, report...
1	 Spotify will now be allowed to offer its own...	spotify will now be allowed to offer its own b...	[spotify, will, now, be, allowed, to, offer, i...	[spotify, allowed, offer, billing, play, store]	[spotify, allow, offer, billing, play, store]
2	 “Buy the dip and stay long on Nvidia” - Jim ...	buy the dip and stay long on nvidia jim cramer	[buy, the, dip, and, stay, long, on, nvidia, j...	[buy, dip, stay, long, nvidia, jim, cramer]	[buy, dip, stay, long, nvidia, jim, cramer]
3	Ryan Cohen sends GameStop shares rallying 	ryan cohen sends gamestop shares rallying	[ryan, cohen, sends, gamestop, shares, rallying]	[ryan, cohen, sends, gamestop, shares, rallying]	[ryan, cohen, send, gamestop, share, rally]
4	 Apple’s services are back up and running ...	apples services are back up and running after ...	[apples, services, are, back, up, and, running...	[apples, services, back, up, running, outage]	[apple, service, back, up, running, outage]
5	 South Korea approves Merck’s Covid pills 	south korea approves mercks covid pills	[south, korea, approves, mercks, covid, pills]	[south, korea, approves, mercks, covid, pills]	[south, korea, approve, merck, covid, pill]
6	Alibaba shares are up nearly 9% after announci...	alibaba shares are up nearly 9 after announcin...	[alibaba, shares, are, up, nearly, 9, after, a...	[alibaba, shares, up, nearly, 9, announcing, r...	[alibaba, share, up, nearly, 9, announce, reco...
7	 Tesla’s German Gigafactory is..	teslas german gigafactory is finally live	[teslas, german, gigafactory, is, finally, live]	[teslas, german, gigafactory, finally, live]	[teslas, german, gigafactory, finally, live]
8	 Nike’s shares are up nearly 6% after reporti...	nikes shares are up nearly 6 after reporting e...	[nikes, shares, are, up, nearly, 6, after, rep...	[nikes, shares, up, nearly, 6, reporting, earn...	[nike, share, up, nearly, 6, reporting, earning]
9	Berkshire Hathaway plans to buy Alleghany ...	berkshire hathaway plans to buy alleghany for ...	[berkshire, hathaway, plans, to, buy, alleghan...	[berkshire, hathaway, plans, buy, alleghany, 1...	[berkshire, hathaway, plan, buy, alleghany...

Tabel 3. Tabel *Software* Dan *Hardware* Pendukung

date	title	stocks	sentiment	rank
24/03/2022	Ouch! Shares of Adobe are down nearly 10%...	ADOBE INC. (XNAS:ADBE)	1	3
24/03/2022	 Spotify will now be allowed to offer its ...	Spotify Technology SA (XNYS:SPOT)	1	3
24/03/2022	 “Buy the dip and stay long on...	NVIDIA CORPORATION (XNAS:NVDA)	0	2
23/03/2022	Ryan Cohen sends GameStop...	GAMESTOP CORP. (XNYS:GME)	1	3
23/03/2022	 Apple’s services are back up and...	APPLE INC. (XNAS:AAPL)	-1	1
23/03/2022	 South Korea approves Merck’s...	MERCK & CO., INC. (XNYS:MRK)	0	2
22/03/2022	Alibaba shares are up nearly 9% after...	Alibaba Group Holding Limited (XNYS:BABA)	1	3
22/03/2022	 Tesla’s German Gigafactory...	TESLA, INC. (XNAS:TSLA)	0	2
22/03/2022	 Nike’s shares are up nearly...	NIKE, INC. (XNYS:NKE)	1	3
21/03/2022	Berkshire Hathaway plans to buy...	BERKSHIRE HATHAWAY INC. (XNYS:BRK.B)	0	2

4.2 Uji Coba Korelasi

Kekuatan korelasi antara nilai sentimen terhadap perubahan harga saham diuji, untuk mencari korelasi terkuat. Ada 4 skenario yang diuji coba:

1) Skenario 1: Korelasi

Pengujian skenario ini menerapkan baik korelasi *spearman* dan *pearson* terhadap nilai sentimen dan perubahan harga saham. Kedua skenario diuji untuk dibandingkan mana nilai yang lebih tinggi. Skenario ini mengimplementasikan tahap-tahap yang sudah dijelaskan sebelumnya, tanpa perubahan apa pun. Hasil dari skenario ini adalah sebagai berikut:

- a. Korelasi Pearson
 - opening1 Correlation: -0.0331
 - closing1 Correlation: 0.0128
 - opening2 Correlation: -0.0024
 - closing2 Correlation: -0.0088
- b. Korelasi Spearman
 - opening1 Correlation: -0.0009
 - closing1 Correlation: 0.0099
 - opening2 Correlation: -0.0123
 - closing2 Correlation: -0.0284

2) Skenario 2: Perubahan Kategorisasi

Pengujian skenario ini mengubah kategorisasi sentimen dengan meniadakan sentimen netral, sehingga sentimen hanya ada 2, yaitu sentimen positif dan sentimen negatif. Perubahan kategorisasi sentimen ini dilakukan dengan asumsi sebagai berikut:

- a. Asumsi 1: Sentimen Netral Dianggap Positif

Mengasumsikan sentimen netral, dengan nilai sentimen sama dengan 0 ($n=0$) adalah sentimen positif. Seluruh sentimen netral akan dianggap sebagai sentimen positif. Kemudian dilakukan penghitungan korelasi dengan metode pearson. Hasil dari asumsi ini adalah sebagai berikut:

 - opening1 Correlation: -0.0691
 - closing1 Correlation: -0.0317
 - opening2 Correlation: -0.0345
 - closing2 Correlation: -0.0339
- b. Asumsi 2: Sentimen Netral Dianggap Negatif

Mengasumsikan sentimen netral, dengan nilai sentimen sama dengan 0 ($n=0$) adalah sentimen negatif. Seluruh sentimen netral akan dianggap sebagai sentimen negatif. Kemudian dilakukan penghitungan korelasi dengan metode pearson. Hasil dari asumsi ini adalah sebagai berikut:

 - opening1 Correlation: 0.0073
 - closing1 Correlation: 0.0452
 - opening2 Correlation: 0.0248
 - closing2 Correlation: 0.0146

3) Skenario 3: Perubahan Saham

Pengujian skenario ini mengubah beberapa saham perusahaan pada data berita yang memuat beberapa saham perusahaan yang berbeda. Skenario ini dilakukan atas dasar bahwa satu berita yang memuat lebih dari satu perusahaan dapat memberikan dampak yang berbeda terhadap perubahan nilai saham perusahaan-perusahaan tersebut. Perubahan ini akan mengubah 50 saham perusahaan yang dimuat oleh 50 data berita. Kemudian dilakukan penghitungan korelasi dengan metode pearson. Hasil dari skenario ini sebagai berikut:

opening1 Correlation: 0.0028
 closing1 Correlation: 0.0334
 opening2 Correlation: 0.0285
 closing2 Correlation: 0.0105

4) Skenario 4: Korelasi Dengan Tidak Mengategorisasikan Nilai Sentimen

Pengujian skenario ini nilai sentimen akan dikorelasikan langsung dengan perubahan nilai saham. Nilai sentimen tidak akan dikategorisasikan ke dalam sentimen positif, negatif, dan netral, melainkan skor korelasi yang ada pada kolom *combined_sentiment* digunakan sebagai variabel korelasi. Jika korelasi negatif bernilai -0,2934, maka nilai -0,2934 itu yang akan menjadi variabel korelasi. Skenario ini akan menggunakan hanya korelasi pearson. Hasil dari skenario ini adalah sebagai berikut:

opening1 Correlation: -0,0377
 closing1 Correlation: -0,0131
 opening2 Correlation: -0,0264
 closing2 Correlation: -0.0213

5. Kesimpulan

5.1 Simpulan

Berdasarkan hasil penelitian yang telah dilakukan, ini diambil beberapa kesimpulan sebagai berikut:

- 1) Berita saham yang dimuat di *newsletter Gotrade Daily* lebih mempengaruhi secara bertolak belakang terhadap perubahan harga saham di hari pertama pembukaan, dibuktikan dengan nilai korelasi yang paling tinggi (0,072) dan konsisten negatif pada seluruh uji coba skenario.
- 2) Berita saham yang dimuat di *newsletter Gotrade Daily* memiliki nilai korelasi yang sangat rendah (hampir mendekati 0) terhadap perubahan harga saham di seluruh waktu, dilihat dari nilai korelasi yang seluruhnya lebih rendah dari 0,1 pada setiap uji coba skenario yang telah dilakukan.
- 3) Uji coba skenario yang dilakukan tidak berhasil meningkatkan angka korelasi secara signifikan. Skenario 2 asumsi pertama memberikan peningkatan angka korelasi yang paling tinggi dibandingkan skenario lain, yaitu pada pembukaan hari pertama (-0.072), namun angka korelasi masih sangat rendah.
- 4) Nilai korelasi yang rendah antara berita saham dengan perubahan nilai saham, membuat berita yang dimuat pada *newsletter Gotrade Daily* tidak dapat diandalkan sebagai satu-satunya pertimbangan untuk mengambil keputusan untuk berinvestasi ke saham tertentu.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, ada beberapa saran yang dapat dipertimbangkan untuk penelitian berikutnya sebagai berikut:

- 1) Ada beberapa faktor eksternal yang mempengaruhi perubahan harga saham, seperti, harga minyak dunia, harga emas, suku bunga *The Fed*, atau isu-isu yang tidak termuat pada *newsletter Gotrade Daily*.
- 2) Analisis sentimen pada penelitian ini hanya menganalisis judul berita saja. Analisis sentimen pada isi berita untuk penelitian selanjutnya dapat memberikan nilai sentimen yang lebih representatif terhadap berita tersebut.

6. Daftar Rujukan

- [1] “Jumlah Investor Saham di Indonesia Lampau 6 Juta SID.” Accessed: Feb. 13, 2025. [Online]. Available: <https://www.idx.co.id/id/berita/siaran-pers/2224>
- [2] R. J. Teweles and E. S. Bradley, *The stock market*, vol. 64. John Wiley & Sons, 1998.
- [3] J. Salim, *Cara gampang bermain saham*. Visimedia, 2010.
- [4] W. Wang, K.-Y. Ho, W.-M. R. Liu, and K. T. Wang, “The relation between news and stock price jump: an analysis based on neural network,” 2013.
- [5] B. Liu, *Sentiment Analysis and Opinion Mining*. in *Synthesis Lectures on Human Language Technologies*. Cham: Springer International Publishing, 2012. doi: 10.1007/978-3-031-02145-9.
- [6] A. D. ’ Andrea, F. Ferri, and P. Grifoni, “Approaches, Tools and Applications for Sentiment Analysis Implementation,” *Int J Comput Appl*, vol. 125, no. 3, pp. 975–8887, 2015, Accessed: Feb. 02, 2025. [Online]. Available: <http://messenger.yahoo.com/features/emoticons>
- [7] A. Alamsyah, S. P. Ayu, and B. Rikumahu, “Exploring relationship between headline news sentiment and stock return,” 2019 7th International Conference on Information and Communication Technology, ICoICT 2019, Jul. 2019, doi: 10.1109/ICOICT.2019.8835298.
- [8] F. Az-Zahra, A. Alamsyah, N. Trianasari, and M. R. Barokah, “The Correlation Analysis of Stock Price Towards Public Sentiment in Automotive Industry,” in 2021 International Conference on Data Science and Its Applications, ICoDSA 2021, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 257–262. doi: 10.1109/ICoDSA53588.2021.9617478.
- [9] A. Rai, “Analisis Sentimen Pemberitaan Vaksin Covid-19 dan Kaitannya dengan Perubahan Harga Saham Emiten Farmasi,” *JURNAL BISNIS STRATEGI*, vol. 30, no. 1, pp. 26–34, Jul. 2021, doi: 10.14710/JBS.30.1.26-34.
- [10] P. M. Nadkarni, L. Ohno-Machado, and W. W. Chapman, “Natural language processing: An introduction,” Mar. 2011. doi: 10.1136/amiajnl-2011-000464.
- [11] E. D. Liddy, “Natural Language Processing,” 2001. [Online]. Available: <https://surface.syr.edu/istpub>
- [12] M. Anandarajan, C. Hill, and T. Nolan, “Text Preprocessing,” pp. 45–59, 2019, doi: 10.1007/978-3-319-95663-3_4.
- [13] K. Smelyakov, D. Karachevtsev, D. Kulemza, Y. Samoilenko, O. Patlan, and A. Chupryna, “Effectiveness of Preprocessing Algorithms for Natural Language Processing Applications,” in 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), 2020, pp. 187–191. doi: 10.1109/PICST51311.2020.9467919.
- [14] Y. Shi, L. Zhu, W. Li, K. Guo, and Y. Zheng, “Survey on Classic and Latest Textual Sentiment Analysis Articles and Techniques,” *Int J Inf Technol Decis Mak*, vol. 18, no. 4, pp. 1243–1287, Mar. 2019, doi: 10.1142/S0219622019300015.
- [15] C. J. Hutto and E. Gilbert, “VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text,” 2014. [Online]. Available: <http://sentic.net/>
- [16] T. Loughran et al., “When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks,” 2011.
- [17] G. Nath, S. Wilson, A. Sood, S. Kavya Durbaka, A. Khanna, and K. Manot, “On Quantifying Sentiments of Financial News Are We Doing the Right Things?,”
- [18] J. L. Rodgers and ; W Alan Nicewander, “Thirteen Ways to Look at the Correlation Coefficient,” 1988.
- [19] S. Boslaugh, “Statistics in a Nutshell (In a Nutshell (O’Reilly)),” no. January 2013, p. 594, 2012. [Online]. Available: https://books.google.com/books/about/Statistics_in_a_Nutshell.html?id=HZpoDjtKT0IC
- [20] Zvi. Bodie et al., “Investments,” p. 2, 2014.
- [21] J. Hartono, “Teori Portofolio dan Analisis Investasi ed.11,” Yogyakarta: BPFE, vol. 470 hlm. :, p. 762, 2017, Accessed: Feb. 07, 2025. [Online]. Available: <https://library.bpk.go.id/koleksi/detil/jkpkbpkpp-p-NokRqwa4W2>
- [22] C. Shu, “Gotrade, the app that lets international users buy fracGotrade, the app that lets international users buy fractions of US stocks, raises \$15.5M Series A,” TechCrunch, 2022.
- [23] Cuevas, G., et.al. “Global Online Trading System (GoTrade),” IEEE Symposium on Systems and Information Engineering Design, SIEDS, 2007.

