

CORRELATION ANALYSIS OF SENTIMENT OF 2024 ELECTION RESULTS AND STOCK MOVEMENTS OF POLITICAL ACTORS IN INDONESIA

Enjelita Sari¹, Lasmedi Afuan², Ipung Permadi³, Eddy Maryanto⁴, Swahesti Puspita Rahayu⁵

^{1,2,3,4,5}Informatics, Engineering Faculty, Universitas Jenderal Soedirman, Indonesia

Email: ¹enjelitasari5@gmail.com, ²lasmedi.afuan@unsoed.ac.id, ³ipung.permadi@unsoed.ac.id,
⁴eddy.maryanto@unsoed.ac.id, ⁵swahesti.rahayu@unsoed.ac.id

(Article received: August 22, 2024; Revision: August 23, 2024; published: August 27, 2024)

Abstract

General elections (elections) are one of the crucial moments in the political life of a country, where the public democratically elects leaders and their deputies to manage the government. Public sentiment towards the results of elections significantly impacts the political stability and economic conditions of a country. This research aims to analyze the relationship between public sentiment towards the 2024 General Elections in Indonesia and changes in the stock prices of political actors using technological approaches and data analysis. The Long Short-Term Memory (LSTM) method is used to classify sentiment based on Twitter data collected with Harvest Tweet. Evaluation of the LSTM model shows an accuracy rate of 90%, precision of 93.6%, and recall of 92.7%. The correlation analysis using the Spearman coefficient indicates a significant negative relationship with a coefficient of 0.402 and a p-value of 0.046. Implementation of an interactive dashboard using Streamlit facilitates visualization of the data used in this study. Recommendations include increasing the amount of training data for sentiment models, exploring alternative correlation methods for deeper analysis, and refining the interface and data integration on the dashboard to enhance user experience and analysis accuracy. This research is expected to contribute to understanding the dynamics of public sentiment and its impact on the stock market in the context of Indonesian politics.

Keywords: correlation, elections, sentiment analysis, stocks.

ANALISIS KORELASI SENTIMEN HASIL PEMILU 2024 DAN PERGERAKAN SAHAM PELAKU POLITIK DI INDONESIA

Abstrak

Pemilihan umum (pemilu) merupakan salah satu momen penting dalam kehidupan politik sebuah negara, di mana masyarakat secara demokratis memilih pemimpin dan wakilnya untuk mengelola pemerintahan. Sentimen publik terhadap hasil pemilu memiliki dampak yang signifikan terhadap stabilitas politik dan keadaan ekonomi suatu negara. Penelitian ini bertujuan untuk menganalisis hubungan antara sentimen publik terhadap hasil Pemilu 2024 di Indonesia dengan perubahan harga saham pelaku politik menggunakan pendekatan teknologi dan analisis data. Metode *Long Short-Term Memory* (LSTM) digunakan untuk mengklasifikasikan sentimen berdasarkan data Twitter yang diambil dengan *Harvest Tweet*. Evaluasi model LSTM menunjukkan tingkat akurasi sebesar 90%, presisi 93,6%, dan *recall* 92,7%. Hasil analisis korelasi menggunakan koefisien *Spearman* menunjukkan hubungan negatif yang signifikan dengan nilai koefisien 0.402 dan nilai p sebesar 0.046. Implementasi *dashboard* interaktif menggunakan *Looker Data Studio* memfasilitasi visualisasi data yang digunakan dalam penelitian ini. Saran yang diberikan termasuk peningkatan jumlah data latih untuk model sentimen, eksplorasi metode korelasi alternatif untuk analisis lebih mendalam, serta penyempurnaan antarmuka dan integrasi data pada *dashboard* untuk meningkatkan pengalaman pengguna dan ketepatan analisis. Penelitian ini diharapkan memberikan kontribusi dalam memahami dinamika sentimen publik dan dampaknya terhadap pasar saham dalam konteks politik Indonesia.

Kata kunci: analisis sentimen, korelasi, pemilu, saham.

1. PENDAHULUAN

Pemilihan umum (pemilu) merupakan peristiwa krusial dalam kehidupan demokrasi suatu negara. Di

Indonesia, pemilu diadakan secara berkala untuk memilih perwakilan rakyat serta pemimpin di tingkat nasional dan daerah. Dengan berbagai isu sosial dan

politik yang menjadi perhatian masyarakat, pemilu Indonesia 2024 menjadi arena politik yang penuh tantangan dan dinamika.

Peristiwa politik, seperti hasil pemilihan, menarik perhatian dunia keuangan dan pasar saham. Keputusan politik yang dihasilkan dari pemilihan dapat mempengaruhi kebijakan pemerintah, stabilitas ekonomi, dan berbagai sektor industri, termasuk perusahaan-perusahaan yang sahamnya diperdagangkan di pasar saham [1]. Menurut CNBC Indonesia pada perdagangan Kamis, 15 Februari 2024 Indeks Harga Saham Gabungan (IHSG) mengalami lonjakan signifikan pada sesi awal perdagangan. Indeks dibuka dengan kenaikan 2%, meskipun kemudian mengalami koreksi dan ditutup dengan kenaikan 1,30% pada level 7.303,28. Lonjakan IHSG ini sebagian besar didorong oleh keyakinan investor terhadap hasil pemilihan presiden 2024 yang dianggap berpotensi berlangsung dalam satu putaran [2].

Berbagai metode yang digunakan untuk merefleksikan pendapat dan persepsi publik mencakup survei opini, liputan media, dan platform media sosial. Salah satu pendekatan untuk memperoleh pemahaman yang lebih mendalam tentang dampak persepsi masyarakat terhadap pilihan politik terhadap pasar keuangan adalah dengan menganalisis korelasi antara sentimen tersebut dengan pergerakan saham tokoh politik.

Penelitian ini berpotensi memberikan kontribusi signifikan terhadap pemahaman hubungan antara ekonomi dan politik di Indonesia. Berdasarkan data Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), pada tahun 2024, sebanyak 79,5% penduduk Indonesia telah terhubung dengan internet [3]. Penelitian ini juga dapat menawarkan wawasan mengenai bagaimana dinamika politik mempengaruhi stabilitas pasar saham yang bermanfaat bagi investor, analisis keuangan, dan pengambilan kebijakan.

Selain memahami hubungan antara ekonomi dan politik di Indonesia, analisis sentimen juga menjadi aspek krusial dalam penelitian ini. Menurut tinjauan literatur oleh Yanying Mao (2024), berbagai metode analisis sentimen yang telah diterapkan dalam penelitian sebelumnya dibandingkan dan tantangan yang dihadapi juga dievaluasi. Dalam literatur tersebut disebutkan bahwa metode *Decision Tree* dan KNN (K-Nearest Neighbor) mudah dipilih dan memiliki waktu pelatihan yang singkat. Namun, kedua metode ini masih sulit diinterpretasikan karena tidak menyediakan penjelasan tentang bagaimana mereka mencapai hasilnya [4].

Untuk mengatasi kelemahan metode tersebut, penelitian lain dilakukan oleh Firdaus (2024), mereka melakukan uji coba analisis sentimen pada *dataset* bakal calon presiden 2024 menggunakan metode SVM (*Support Vector Machine*) dan *Naïve Bayes*, yang dapat memberikan interpretasi hasil lebih baik dibandingkan dengan *Decision Tree* dan KNN.

Penelitian ini mencapai nilai akurasi sebesar 70% hingga 80% [5]. Namun, metode SVM dan *Naïve Bayes* yang digunakan oleh Firdaus masih memiliki kelemahan, yaitu tidak mampu menangkap ketergantungan jangka panjang dalam teks. Artinya, mereka tidak dapat mempertimbangkan konteks keseluruhan kalimat saat menganalisis sentimen.

Untuk mengatasi keterbatasan metode seperti *Decision Tree* dan KNN yang kurang dapat diinterpretasikan dengan baik, serta SVM dan *Naïve Bayes* yang tidak mampu menangkap ketergantungan jangka panjang dalam teks, beberapa penelitian telah mengadopsi metode GRU (*Gated Recurrent Unit*) untuk analisis sentimen. Misalnya, penelitian oleh Abdelgwad (2022) menunjukkan bahwa struktur jaringan GRU dan LSTM (*Long Short-Term Memory*) memiliki kesamaan dan memberikan hasil yang hampir setara. Kelebihan GRU terletak pada kesederhanaan model dan kecepatan konvergensi yang lebih tinggi. Namun, dalam kasus data berukuran besar, LSTM menunjukkan kinerja yang lebih baik karena memiliki lebih banyak gerbang dan parameter [6].

Dalam penelitian ini, LSTM dipilih sebagai metode analisis sentimen karena kemampuannya untuk mengatasi keterbatasan dari beberapa metode yang telah disebutkan sebelumnya. LSTM efektif dalam menangkap ketergantungan jangka panjang, memberikan interpretasi hasil yang lebih baik, dan dapat diterapkan pada *dataset* besar. Pemilihan metode LSTM didukung oleh penelitian terdahulu yang dilakukan oleh Nurkholis (2022) dan Romadhoni (2022). Kedua penelitian tersebut menunjukkan bahwa LSTM memiliki tingkat akurasi yang lebih tinggi dibandingkan dengan *Naïve Bayes*, dengan nilai akurasi LSTM sebesar 77% dan presisi sebesar 84%, sedangkan nilai akurasi *Naïve Bayes* hanya sebesar 76% dengan presisi sebesar 75% [7], [8]. LSTM memungkinkan model untuk mempertimbangkan konteks kalimat secara keseluruhan, yang sangat krusial dalam analisis sentimen, dan juga memungkinkan visualisasi bobot jaringan saraf untuk pemahaman yang lebih mendalam mengenai cara model mencapai hasilnya. Meskipun LSTM memiliki kelemahan berupa kompleksitas model dan waktu pelatihan yang tinggi, kelemahan ini dapat diatasi dengan metode GRU. Namun, pada *dataset* yang besar, LSTM tetap lebih unggul karena memiliki lebih banyak gerbang dan parameter, yang menghasilkan hasil yang lebih akurat dan komprehensif [6].

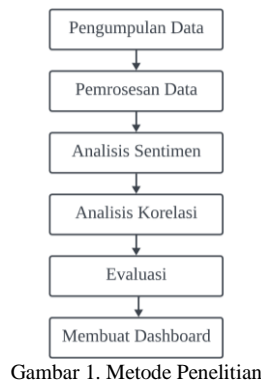
Selanjutnya, penelitian ini juga mengeksplorasi korelasi antara sentimen publik terhadap hasil pemilu dan pergerakan saham tokoh politik. Korelasi yang dianalisis dalam penelitian ini merujuk pada hubungan statistik antara dua variabel: sentimen publik terhadap hasil pemilu yang diukur melalui analisis teks media sosial, dan pergerakan saham tokoh politik yang diukur dari perubahan harga saham di pasar modal. Korelasi *Spearman* digunakan untuk

menilai hubungan tersebut dalam konteks statistik non-parametrik atau skala ordinal [9]. Korelasi ini merupakan ukuran korelasi antara dua variabel yang diukur minimal dalam skala ordinal sehingga objek penelitian dapat diurutkan dalam dua rangkaian yang berurutan [10].

Oleh karena itu, diharapkan bahwa penelitian mengenai pengaruh antara hasil Pemilu 2024 dan pergerakan saham tokoh politik di Indonesia dapat memperdalam pemahaman mengenai hubungan kompleks antara elemen ekonomi dan politik dalam demokrasi modern.

2. METODE PENELITIAN

Metode yang digunakan dalam penelitian ini ditunjukkan pada gambar 1. Metode dilakukan secara bertahap dan berurutan.



Gambar 1. Metode Penelitian

Seperti yang tertera pada gambar 1, metode penelitian dilakukan secara bertahap mulai dari pengumpulan data sampai pembuatan dashboard. Penjelasan lebih lanjut mengenai metode penelitian yang dilakukan adalah sebagai berikut.

2.1. Pengumpulan Data

Pengumpulan data merupakan tahapan penelitian berupa mengumpulkan data yang akan digunakan untuk analisis dan dikorelasikan. Data dibagi menjadi tiga bagian, yaitu sebagai berikut.

2.1.1. Pengumpulan Data Latih

Data latih, yang berisi teks sentimen publik dengan label positif maupun negatif, digunakan untuk melatih model agar terbiasa dengan konteks pembahasan politik. Contoh data latih yang akan digunakan untuk melatih model ditunjukkan pada tabel 1.

Teks	Label
Kampanye pemilu kali ini sangat informatif dan transparan, saya sangat terkesan dengan program-program yang ditawarkan	Positive
Banyak janji-janji palsu dari calon-calon pemilu, saya tidak percaya mereka akan menepati apa yang mereka katakan.	Negative
Pelaksanaan pemilu berjalan lancar dan tertib, petugas sangat membantu dan profesional.	Positive

Tabel 1 berisi kumpulan kalimat (teks) dan label sentimennya (positif atau negatif). Data ini akan digunakan untuk melatih model komputer agar bisa secara otomatis menentukan apakah suatu kalimat itu mengandung sentimen positif atau negatif. Dengan kata lain, model akan belajar dari contoh-contoh yang diberikan dalam tabel untuk kemudian memprediksi sentimen dari kalimat baru yang belum pernah dilihat sebelumnya.

2.1.2. Pengumpulan Data Uji

Data uji, yang berupa data hasil *crawling* dari media sosial yang berisikan *post* pendapat pengguna terhadap hasil pemilu. Contoh data sentimen yang dikumpulkan sebagai data uji ditunjukkan pada tabel 2.

created_at	full_text	location	user_id	username
Mon Mar 25 2024	Pemilu kali ini benar2 parah, banyak kecurangan dimana2! #pemilu #curang 😞😞	Yogyakarta	21354648	usn_1
Mon Mar 25 2024	Hari ini ke TPS buat nyoblos, semoga hasilnya nggak mengecewakan 🙏 #pemilu	Bandung	856515123	usn_2
Mon Mar 25 2024	Gimana pendapat kalian soal debat capres tadi malam? Seru banget kan!! #debatCapres #pemilu2024	Depok	455545214	usn_3

Tabel 2 berisi kumpulan komentar atau opini publik tentang pemilu. Setiap baris mewakili satu komentar lengkap dengan informasi siapa yang menulisnya, kapan, dan dari mana. Data ini bisa digunakan untuk menganalisis sentimen masyarakat terhadap pemilu, topik yang sedang ramai dibicarakan, dan perbedaan pendapat antar daerah.

2.1.3. Pengumpulan Data Saham

Data pasar saham, dibutuhkan untuk nantinya dikorelasikan dengan hasil analisis sentimen. Contoh data pasar saham ditunjukkan pada tabel 3.

Tanggal	Terakhir	Pembukaan	Tertinggi	Terendah	Vol.	Perubahan%
28/04/2024	1.020	985	1.095	985	104,30K	3,55%
21/04/2024	985	1.015	1.040	980	71,80K	-2,48%
14/04/2024	1.010	1.025	1.260	1000	256,90K	-1,46%
31/03/2024	1.020	1.030	1.030	1.005	112,60K	0,49%
24/03/2024	1.020	1.000	1.080	970	130,10K	2,51%

Tabel 3 menampilkan data historis harga suatu aset finansial dalam beberapa hari terakhir. Setiap baris merepresentasikan data harga pada satu hari tertentu, meliputi harga pembukaan, penutupan, tertinggi, terendah, volume perdagangan, serta perubahan persentase harga dibandingkan hari sebelumnya. Data ini berguna untuk menganalisis pergerakan harga, volatilitas, dan tren pasar.

2.2. Prapemrosesan Data

Prapemrosesan data adalah tahap untuk mempersiapkan data agar dapat diolah dan hasil yang

didapatkan menjadi lebih baik dan akurat. Pemrosesan data melalui beberapa tahapan sebagai berikut.

2.2.1. *Translating*

Proses *translating* dilakukan karena data latih yang digunakan berbahasa Inggris. Dengan demikian, data yang akan di analisis dapat digunakan secara konsisten dengan *dataset* tersebut dan memungkinkan analisis dan pelatihan model yang akurat. Contoh hasil dari proses *translating* ditunjukkan pada tabel 4.

Tabel 4. Contoh Hasil *Translating*

Sebelum <i>Translating</i>	Setelah <i>Translating</i>
Pemilu kali ini benar2 parah, banyak kecurangan dimana2! #pemilu #curang 😞😞	This election is really bad, there is a lot of cheating everywhere! #election #cheating 😞😞
Hari ini ke TPS buat nyoblos, semoga hasilnya nggak mengecewakan 🙏 #pemilu	Going to the polling station today to vote, hopefully the results will not be disappointing 🙏 #election
Gimana pendapat kalian soal debat capres tadi malam? Seru banget kan!! #debatCapres #pemilu2024	What do you think about the presidential debate last night? It was really exciting, right!! #presidentialdebate #election2024

Tabel 4 menunjukkan contoh penerjemahan kalimat dari bahasa Indonesia ke bahasa Inggris. Kalimat-kalimat aslinya diambil dari media sosial dan berkaitan dengan pemilu. Terjemahannya cukup akurat dan mencerminkan nuansa bahasa sehari-hari.

2.2.2. *Cleansing*

Pada tahapan ini dilakukan penghapusan data teks yang sama atau redundan dan menghapus simbol yang tidak diperlukan. Contoh dari hasil proses *cleansing* seperti yang ada pada tabel 5.

Tabel 5. Contoh Hasil *Cleansing*

Sebelum <i>Cleansing</i>	Setelah <i>Cleansing</i>
This election is really bad, there is a lot of cheating everywhere! #election #cheating 😞😞	This election is really bad there is a lot of cheating everywhere
Going to the polling station today to vote, hopefully the results will not be disappointing 🙏 #election	Going to the polling station today to vote hopefully the results will not be disappointing
What do you think about the presidential debate last night? It was really exciting, right!! #presidentialdebate #election2024	What do you think about the presidential debate last night It was really exciting right

Tabel 5 menunjukkan hasil dari proses *cleansing* dimana adanya *hashtag*, tanda baca, dan juga emoji telah dihapus dari teks.

2.2.3. *Case folding*

Case folding, yaitu proses penyeragaman terhadap teks dengan mengubah seluruh huruf pada teks menjadi *lowercase*. Contoh hasil pada proses ini adalah seperti pada tabel 6.

Tabel 6 menunjukkan hasil dari proses *case folding* dimana semua huruf kapital pada teks sebelumnya diubah menjadi *lowercase*.

Tabel 6. Contoh Hasil *Case folding*

Sebelum <i>Case Folding</i>	Setelah <i>Case Folding</i>
This election is really bad there is a lot of cheating everywhere	this election is really bad there is a lot of cheating everywhere
Going to the polling station today to vote hopefully the results will not be disappointing	going to the polling station today to vote hopefully the results will not be disappointing
What do you think about the presidential debate last night It was really exciting right	what do you think about the presidential debate last night It was really exciting right

2.2.4. *Filtering*

Pada tahap ini kata-kata yang tidak diperlukan seperti “yang”, “ini”, “kalo”, dan lainnya akan dihapus dan hanya kata-kata bermakna yang akan digunakan untuk analisis lebih lanjut. Contoh hasil dari *Filtering* ditunjukkan pada tabel 7.

Tabel 7. Contoh Hasil *Filtering*

Sebelum <i>Filtering</i>	Setelah <i>Filtering</i>
this election is really bad there is a lot of cheating everywhere	election really bad lot cheating everywhere
going to the polling station today to vote hopefully the results will not be disappointing	going polling station today vote hopefully results disappointing
what do you think about the presidential debate last night It was really exciting right	think presidential debate last night really exciting right

Tabel 7 menunjukkan contoh hasil dari proses *filtering*, dimana kata-kata yang tidak memiliki makna akan dihapus dari teks sehingga hanya tersisa kata-kata yang bisa dijadikan acuan untuk analisis sentimen.

2.2.5. *Tokenizing*

Tokenizing adalah proses dimana data teks dipisahkan tiap kata berdasarkan spasi. Contoh hasil dari proses *tokenizing* ditunjukkan pada tabel 8.

Tabel 8. Contoh Hasil *Tokenizing*

Sebelum <i>Tokenizing</i>	Setelah <i>Tokenizing</i>
election really bad lot cheating everywhere	['election', 'really', 'bad', 'lot', 'cheating', 'everywhere']
going polling station today vote hopefully results disappointing	['going', 'polling', 'station', 'today', 'vote', 'hopefully', 'results', 'disappointing']
think presidential debate last night really exciting right	['think', 'presidential', 'debate', 'last', 'night', 'really', 'exciting', 'right']

Tabel 8 menampilkan contoh hasil dari proses tokenisasi, di mana setiap kata dalam data diubah menjadi sebuah token.

2.2.6. *Stemming*

Pada tahap ini kata yang berimbuhan diubah menjadi kata dasar berdasarkan kamus yang telah disediakan. Contoh hasil dari proses *stemming* ditunjukkan pada tabel 9.

Tabel 9. Contoh Hasil *Stemming*

Sebelum <i>Stemming</i>	Setelah <i>Stemming</i>
['election', 'really', 'bad', 'lot', 'cheating', 'everywhere']	['elect', 'realli', 'bad', 'lot', 'cheat', 'everywher']
['going', 'polling', 'station', 'today', 'vote', 'hopefully', 'results', 'disappointing']	['go', 'poll', 'station', 'today', 'vote', 'hope', 'result', 'disappoint']
['think', 'presidential', 'debate', 'last', 'night', 'really', 'exciting', 'right']	['think', 'presidenti', 'debat', 'last', 'night', 'realli', 'excit', 'right']

Tabel 9 menampilkan hasil dari *stemming*, dimana setiap kata diubah menjadi bentuk kata dasarnya.

2.3. Analisis Sentimen

Analisis sentimen adalah tahap untuk mengekstrak dan menganalisis opini, emosi, dan sentimen dari data teks. Ini melibatkan penggunaan berbagai teknik NLP (*Natural Language Processing*) untuk memahami makna dan konteks teks, dan kemudian mengklasifikasikannya berdasarkan sentimen positif dan negatif [11]. Tahap-tahap analisis sentimen yang akan dilakukan adalah sebagai berikut.

2.3.1. Mempersiapkan Data Latih

Pada tahap ini data label teks yang mewakili kategori atau sentimen diubah menjadi representasi numerik dengan menggunakan teknik *LabelEncoder*. Hal ini dilakukan karena algoritma pembelajaran mesin umumnya bekerja lebih baik dengan data numerik. Selanjutnya data dibagi menjadi dua kumpulan set yaitu set pelatihan dan set pengujian. Pembagian ini dilakukan dengan menggunakan fungsi *train_test_split* dari pustaka *scikit-learn*. Contoh *encoding* pelabelan adalah seperti yang ada pada tabel 10.

Tabel 10. Contoh Hasil *Encoding*

Kategori	Encoding
<i>Positive</i>	1
<i>Negative</i>	0

Tabel 10 menunjukkan representasi numerik dari teks dengan cara mengonversi setiap kata atau token menjadi angka. Setiap kata dalam teks diberikan nomor identifikasi yang unik.

2.3.2. Set *Hyperparameter*

Hyperparameter adalah parameter yang menentukan bagaimana model dilatih dan dioptimalkan. Beberapa parameter yang diset adalah *vocab_size*, *oov_tok*, *embedding_dim*, *max_length*, *padding_type*, dan *trunc_type*.

2.3.3. *Tokenizing* dan *Padding*

Tahap ini dilakukan untuk memastikan semua kalimat pelatihan diubah menjadi string, membuat *tokenizer* dan fit pada kalimat pelatihan. Selain itu pada tahap ini dilakukan konversi kalimat pelatihan menjadi urutan angka dengan *sequences* dan juga dilakukan *padding* dengan menambahkan token khusus (biasanya token 0) ke urutan yang lebih pendek. Hal ini dilakukan untuk memastikan bahwa model dapat memproses semua kalimat dengan cara yang sama, terlepas dari panjangnya. Kemudian langkah yang sama diterapkan pada data pengujian.

2.3.4. Membuat Model LSTM

Beberapa langkah yang dilakukan pada tahap ini adalah mendefinisikan model, menambahkan lapisan *embedding*, menambahkan lapisan LSTM *Bidirectional*, dan menambahkan lapisan *dense* [12], [13].

2.3.5. Kompilasi dan Ringkasan Model

Tahap ini melibatkan menentukan pengoptimal, fungsi kerugian, dan metrik yang akan digunakan selama pelatihan.

2.3.6. Pelatihan Model Menggunakan Data Latih

Tahap ini melibatkan proses memasukkan urutan *input* dan label yang sesuai ke model, dan memungkinkan model untuk menyesuaikan bobotnya untuk meminimalkan fungsi kerugian.

2.3.7. Implementasi Model pada Data Uji

Setelah model dilatih dengan beberapa kali *epoch*, langkah terakhir yang dilakukan adalah menerapkan model untuk memprediksi sentimen pada teks hasil pemilu yang telah dikumpulkan. Contoh hasil dari analisis sentimen yang diterapkan ditunjukkan pada tabel 11.

Tabel 11. Contoh Hasil *Tokenizing*

Teks	Hasil Analisis Sentimen
['elect', 'realli', 'bad', 'lot', 'cheat', 'everywher']	<i>Negatif</i>
['go', 'poll', 'station', 'today', 'vote', 'hope', 'result', 'disappoint']	<i>Positif</i>
['think', 'presidenti', 'debat', 'last', 'night', 'realli', 'excit', 'right']	<i>Positif</i>

Tabel 11 menunjukkan hasil analisis dari data teks yang telah dibersihkan, menghasilkan beberapa klasifikasi, yaitu positif dan negatif.

2.4. Analisis Korelasi

Analisis korelasi adalah metode statistik yang digunakan untuk mengukur hubungan antara dua variabel. Ini membantu untuk memahami apakah ada hubungan antara dua variabel, dan seberapa kuat hubungan tersebut [14], [15]. Misalnya hasil korelasi menunjukkan:

- Korelasi antara sentimen positif dan perubahan harga saham: 0.85
- Korelasi antara sentimen negatif dan perubahan harga saham: -0.65

Dari hasil tersebut dapat diinterpretasikan menjadi:

- Positif (+0.85) menunjukkan adanya korelasi positif yang kuat antara sentimen positif *tweet* dan kenaikan harga saham. Artinya, semakin banyak *tweet* dengan sentimen positif, harga saham cenderung meningkat.
- Negatif (-0.65) menunjukkan adanya korelasi negatif yang cukup kuat antara sentimen negatif *tweet* dan penurunan harga saham. Artinya,

semakin banyak *tweet* dengan sentimen negatif, harga saham cenderung menurun.

2.5. Evaluasi

Setelah menyelesaikan analisis sentimen dan analisis korelasi, tahap selanjutnya adalah evaluasi. Evaluasi ini penting untuk memastikan bahwa hasil analisis valid dan dapat diandalkan.

2.5.1. Evaluasi Analisis Sentimen

Dalam bidang machine learning, *confusion matrix* adalah tabel khusus yang diadopsi untuk mendeskripsikan dan menilai kinerja model klasifikasi untuk sekumpulan data uji yang fitur pembedanya telah diketahui atau supervised learning [16]. *Confusion matrix* terdiri atas 2x2 matriks dengan besaran nilai biner bernilai 1 atau 0. Dengan besaran tersebut dapat diperoleh 4 kemungkinan keluaran yang didefinisikan pada tabel. Untuk penjelasan lebih lanjut bisa dilihat pada tabel 12.

Tabel 12. *Confusion matrix*

		Realita	
		Positive	Negative
Prediksi	True	True Positive	True Negative
	False	False Positive	False Negative

Berdasarkan tabel 12 yang merupakan *confusion matrix*, parameter yang akan digunakan untuk mengukur validitas adalah sebagai berikut:

1. Akurasi, adalah kedekatan nilai prediksi dengan nilai aktual. Rumus perhitungan akurasi adalah pada persamaan 1.

$$Akurasi = \frac{True\ Positives + True\ Negatives}{True\ Positives + False\ Positives + True\ Negatives + False\ Negatives} \quad (1)$$

2. *Precision*, adalah tingkat ketepatan sistem untuk menentukan data yang relevan antara data yang diuji dan data yang dihasilkan dari sistem. Rumus untuk perhitungan *precision* adalah persamaan 2.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (2)$$

3. *Recall*, adalah rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar. Rumus perhitungan *recall* adalah persamaan 3.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (3)$$

2.6. Dashboard

Langkah selanjutnya yang dilakukan adalah membuat *dashboard* interaktif sehingga data dan hasil dari analisis dapat tersampaikan dan mudah dipahami oleh audiens. *Dashboard* dibuat dengan menggunakan *tools Looker studio* dengan mempertimbangkan desain *dashboard*, visualisasi

data termasuk grafik apa yang akan digunakan, dan menambahkan elemen interaktivitas seperti filter, *dropdown* sehingga pengguna dapat menyesuaikan tampilan sesuai dengan kebutuhan, hal ini bertujuan untuk meningkatkan pengalaman pengguna dan memungkinkan untuk dapat mengeksplorasi data lebih dalam [17].

2.6.1. Use case

Pada tahap ini dilakukan identifikasi dan definisi kebutuhan sistem berdasarkan interaksi antara pengguna (aktor) dan sistem [18].

2.6.2. Perancangan

Pada tahap ini, dibuat *mockup dashboard* untuk memberikan gambaran awal mengenai desain dan fungsionalitas yang akan diimplementasikan. Hal ini bertujuan untuk memastikan bahwa pembuatan *dashboard* lebih terstruktur dan terencana dengan baik.

2.6.3. Implementasi

Pada tahap ini dilakukan implementasi dari rancangan atau *mockup* yang telah dibuat sebelumnya.

3. HASIL DAN PEMBAHASAN

Pada bagian ini dapat diuraikan mengenai hasil dan pembahasan dari penelitian yang dilakukan.

3.1. Pengumpulan Data

Pada penelitian ini, pengumpulan data dibagi menjadi tiga bagian, yaitu pengumpulan data yang akan digunakan untuk melatih model, data yang akan diuji, dan data saham untuk dikorelasikan dengan hasil analisis sentimennya.

3.1.1. Pengumpulan Data Latih

Data latih yang digunakan pada penelitian ini diambil dari penelitian yang sebelumnya dilakukan oleh Asno Azzawagama (2024) mengenai pengumpulan *dataset* bakal calon presiden Indonesia 2024. Data latih yang digunakan adalah sebanyak 30.000 data. *Dataset* yang digunakan untuk melatih model ditunjukkan pada tabel 12. Akses penuh ke data uji yang digunakan dalam penelitian ini dapat diperoleh melalui tautan <https://data.mendeley.com/datasets/7w5zvr8jgp/5>.

Tabel 12. Data Latih Penelitian

index	Date	User ID	tweetcontent	Text	label
0	2023-04-16 09:44:01+00:00	1.02E+18	ibairil Anwar	gemira and semeton ball muslims declare support for prabowo as president. news updates	Positive
1	2023-04-16 09:27:10+00:00	1.19E+18		fahnoor inismynname that's why we support the president	Negative
2	2023-04-16 08:32:06+00:00	1.60E+18		this time, jokowi's quota for wrinkles on his face and white hair is identical to his aura, sir, moving on to answering sincerely and honestly, agree, that's great, prabowo subianto	Positive
3	2023-04-16 08:19:16+00:00	4745573953		happy birthday to kopassus, may the spirit of patriotism continue to burn in our hearts, be the front guard in maintaining the sovereignty and security of the republic of indonesia, keep advancing, kopassus, always healthy, mr. general prabowo subianto, president of the republic of indonesia	Positive
...
29998	2023-03-05 07:03:25+00:00	1.61E+18		and various ethnicities are here, there are lampung, there are javanese, there are padang, madura, there are bugis. we are united in this place because we hope for our president, who asked muzani to the thousands of attendees prabowo answered the audience	Positive
29999	2023-03-05 06:47:22+00:00	1.52E+18		look at the people smiling when they meet pak prabowo, hopefully pak prabowo's dream of making the little people too gemuyu will be granted and made easier in the hands of pak prabowo's leadership	Positive

Berdasarkan tabel 12, data latih yang digunakan mencakup berbagai informasi seperti tanggal pembuatan *tweet*, *User ID*, lokasi, konten *tweet*, serta label sentimen.

3.1.2. Pengumpulan Data Uji

Pengumpulan data dilakukan menggunakan *Google Colaboratory* menggunakan bantuan modul bernama *Tweet-Harvest*, pemilihan modul ini dikarenakan kemudahan dan pengambilan datanya yang dapat memuat banyak data. Jumlah data *tweet* yang dikumpulkan adalah 100 data per minggu dari setiap kata kunci, dengan total pengumpulan data selama 5 minggu dan menggunakan 7 kata kunci berbeda. Jika dihitung, total data yang terkumpul adalah sebanyak 3.448 *tweet*. Hasil dari data yang telah dikumpulkan melalui proses *crawling* ini ditunjukkan pada tabel 13. Akses penuh ke data uji yang digunakan dalam penelitian ini dapat diakses melalui <https://www.kaggle.com/datasets/enjelita/data-tugas-akhir>.

Tabel 13. Data Uji Penelitian

ID	created_at	full_text	location	user_id_str	username
1	Mon Mar 25 23:58:31 +0000 2024	@H4T14K4LN4L42 @prabowo @jokowi @Gerindra @PartaiGolkar @PartaiPAN @Pdemokrat @KPU_ID @bawaslu_RI Kmn mah jadi pengamat skrg udah jadi pakar politik...		77471755	donald29
2	Mon Mar 25 23:57:45 +0000 2024	@H4T14K4LN4L42 @prabowo @jokowi @Gerindra @PartaiGolkar @PartaiPAN @Pdemokrat @KPU_ID @bawaslu_RI Dasar Polisi Bajigur katanya Menang klo Menang hadapi dong Secara GENTLEMEN pertahankan KEMENANGAN 02 lewat adu Argumentasi versi ANGKET... Siap2lah hancur berantakan Koalisi 02 km sdh di kepung dua Jurusan Belakang dan Muka. Hak Angket dan Jalur MK...	Yogyakarta, Indonesia	1.6767E+18	KhairuFah87870
3	Mon Mar 25 23:52:35 +0000 2024	@Encang_ajja Megamendung bknya daerah rnh pak @prabowo ya?tolong min @Gerindra ada warga yg perlu dibantu.	Bdg sebelah sana	1.60575E+18	kabasyank
3446	Mon Apr 29 23:58:46 +0000 2024	@Muhammad_SaeudPDIp sm wniwiy ga beda jauh sm2 pembahung dan nyengsarain rhyt.	Banyuwangi, Jawa Timur	1.71922E+18	Nawrakal0
3447	Mon Apr 29 23:53:52 +0000 2024	Kemapa Pak Prabowo berani ngomong begini? Kecapusan? Pdhi ini bisa jadi alat bukti baru pengakuan dari csara terpbh atas cawad presiden. Karena dia tau semua elemen ga bakal ajukan hak angket selain PKS pdhi cuman butuh min 2 fraksi. PDIP? Hehehe... https://t.co/9x2Jv35QO	Nangro Aceh Darussalam, Indone	2151205796	masawep08

Berdasarkan tabel 13, data uji dalam penelitian ini meliputi beberapa kolom, antara lain tanggal pembuatan *tweet*, isi *tweet*, lokasi, *User ID*, dan *username*.

3.1.3. Pengumpulan Data Saham

Pengumpulan data saham dilakukan dengan menggunakan *Investing.com* yang merupakan platform layanan penyedia informasi terkait pasar saham. Untuk mengumpulkan data saham diperlukan data historis harian dari pasar saham tersebut yang berisikan harga saham pembukaan, penutupan, volume perdagangan, dan indikator kinerja seperti yang ditunjukkan pada tabel 14. Akses penuh ke data saham yang digunakan dalam penelitian ini dapat diperoleh melalui tautan <https://www.kaggle.com/datasets/enjelita/data-tugas-akhir>.

Dari tabel 14 dapat dijelaskan bahwa data historis pasar saham terkumpul sebanyak 500 data. Data tersebut diambil dari data harian dari 20 Maret 2024 sampai 03 Mei 2024 di setiap saham.

Tabel 14. Data Saham

index	Tanggal	Terakhir	Pembukaan	Tertinggi	Terendah	Vol.	Perubahan
0	14/04/2024	1.53	1.49	1.575	1.475	54,62M	2,68%
1	14/04/2024	1.53	1.49	1.575	1.475	54,62M	2,68%
2	14/04/2024	1.53	1.49	1.575	1.475	54,62M	2,68%
3	14/04/2024	1.53	1.49	1.575	1.475	54,62M	2,68%
...
498	31/03/2024	1.49	1.445	1.525	1.425	58,40M	3,11%
499	31/03/2024	1.49	1.445	1.525	1.425	58,40M	3,11%

3.2. Prapemrosesan

Data Twitter yang telah dikumpulkan masih dalam bentuk tidak terstruktur dan memerlukan normalisasi sebelum dapat digunakan untuk analisis. Dari total 3.448 data yang terkumpul, akan diambil 4 data sebagai sampel.

3.2.1. Translating

Untuk menyesuaikan dengan data latih yang berbahasa Inggris, maka data yang akan dianalisis diubah menjadi bahasa Inggris melalui proses *translating*. Hasil dari proses *translating* ditunjukkan pada tabel 15.

Tabel 15. Hasil Proses *Translating*

full_text	translated
@H4T14K4LN4L42 @prabowo @jokowi @Gerindra @PartaiGolkar @PartaiPAN @Pdemokrat @KPU_ID @bawaslu_RI Kmn mah jadi pengamat skrg udah jadi pakar politik...	@H4t14k4ln4l42 @prabowo @jokowi @gerindra @partaigolkar @partaipan @pdemokrat @kpu_id @bawaslu_ri kmn still so the observer now is a political expert politik...
@H4T14K4LN4L42 @prabowo @jokowi @Gerindra @PartaiGolkar @PartaiPAN @Pdemokrat @KPU_ID @bawaslu_RI Dasar Polisi Bajigur katanya Menang klo Menang hadapi dong Secara GENTLEMEN pertahankan KEMENANGAN 02 lewat adu Argumentasi versi ANGKET... Siap2lah hancur berantakan Koalisi 02 km sdh di kepung dua Jurusan Belakang dan Muka. Hak Angket dan Jalur MK...	@H4t14k4ln4l42 @prabowo @jokowi @gerindra @partaigolkar @partaipan @pdemokrat @kpu_id @bawaslu_ri Basic Bajigur politician said he wins if you win facing the Gentelmen.Back and face majors.Questionnaire Rights and MK Line ...
@Encang_ajja Megamendung bknya daerah rnh pak @prabowo ya?tolong min @Gerindra ada warga yg perlu dibantu.	@Encang_ajja megamendung instead of the area of ꦱꦺꦴꦠꦺꦤ꧀ the house sir @prabowo huh? Please min @gerindra there are residents who need to be helped.
@H4T14K4LN4L42 @prabowo @jokowi @Gerindra @PartaiGolkar @PartaiPAN @Pdemokrat @KPU_ID @bawaslu_RI Hanya gimick....itu mereka	@H4t14k4ln4l42 @prabowo @jokowi @gerindra @partaigolkar @partaipan @pdemokrat @kpu_id @bawaslu_ri only gimmick

Tabel 15 menampilkan hasil dari proses translasi, di mana data yang awalnya berbahasa Indonesia diubah menjadi bahasa Inggris untuk menyesuaikan dengan data latih yang digunakan.

3.2.2. Cleansing

Pada tahap *cleansing*, data yang telah diterjemahkan ke bahasa Inggris akan dibersihkan dengan menghilangkan angka, simbol, dan juga emoji. Hasil dari *cleansing* data ditunjukkan pada tabel 16.

Tabel 16. Hasil Proses *Cleansing*

translated	cleaned
@H4t14k4ln4l42 @prabowo @jokowi @gerindra @partaigolkar @partaipan @pdemokrat @kpu_id @bawaslu_ri kmn still so the observer now is a political expert	kmn still so the observer now is a political expert
@H4t14k4ln4l42 @prabowo @jokowi @gerindra @partaigolkar @partaipan @pdemokrat @kpu_id @bawaslu_ri Basic Bajigur politician said he wins if you win facing the Gentelmen.Back and face majors.Questionnaire Rights and MK Line ...	Basic Bajigur politician said he wins if you win facing the Gentelmen Back and face majors Questionnaire Rights and MK Line
@Encang_ajja megamendung instead of the area of ꦱꦺꦴꦠꦺꦤ꧀ the house sir @prabowo huh? Please min @gerindra there are residents who need to be helped.	megamendung instead of the area of the house sirhuh Please minthere are residents who need to be helped
@H4t14k4ln4l42 @prabowo @jokowi @gerindra @partaigolkar @partaipan @pdemokrat @kpu_id @bawaslu_ri only gimmick	only gimmick

Tabel 16 menunjukkan hasil dari proses pembersihan data, di mana teks yang sebelumnya mengandung banyak *tag* akun, *hashtag*, emoji, dan tanda baca telah dihapus, sehingga hanya menyisakan teks yang bersih.

3.2.3. Case folding

Setelah data teks dibersihkan dari angka dan simbol, selanjutnya huruf pada data akan disamakan menjadi *lowercase* semua untuk mempermudah analisis teks nantinya. Pada tabel 17 ditunjukkan hasil dari pengubahan dari *Case folding* yaitu tidak adanya huruf kapital pada data teks.

Tabel 17. Hasil Proses *Case folding*

<i>cleaned</i>	<i>lower</i>
kmrn still so the observer now is a political expert	kmrn still so the observer now is a political expert
Basic Bajigur politician said he wins if you win facing the Gentlemen Back and face majors Questionnaire Rights and MK Line	basic bajigur politician said he wins if you win facing the gentlemen back and face majors questionnaire rights and mk line
megamendung instead of the area of the house sirhuh Please minthere are residents who need to be helped	megamendung instead of the area of the house sirhuh please minthere are residents who need to be helped
only gimmick	only gimmick

Tabel 17 memperlihatkan hasil dari proses *case folding*, di mana teks yang sebelumnya masih mengandung huruf kapital telah diubah seluruhnya menjadi huruf kecil atau *lowercase*.

3.2.4. Filtering

Pada proses *Filtering* dilakukan penghapusan kata-kata yang termasuk dalam daftar *stopwords* dalam bahasa Inggris dan juga daftar tambahan *stopwords* dari file teks untuk mengurangi jumlah data yang tidak relevan dan fokus pada kata-kata yang lebih bermakna untuk analisis sentimen. Hasil dari *Filtering* ditunjukkan pada tabel 18.

Tabel 18. Hasil Proses *Filtering*

<i>lower</i>	<i>filtered</i>
kmrn still so the observer now is a political expert	still observer political expert
basic bajigur politician said he wins if you win facing the gentlemen back and face majors questionnaire rights and mk line	basic bajigur politician said wins win facing gentlemen back face majors questionnaire rights mk line
megamendung instead of the area of the house sirhuh please minthere are residents who need to be helped	megamendung instead area house sirhuh please minthere residents need helped
only gimmick	gimmick

Tabel 18 menunjukkan hasil dari proses *filtering*, di mana teks yang sebelumnya mengandung banyak kata-kata tidak berarti, setelah proses *filtering* hanya menyisakan kata-kata yang memiliki makna dan relevan untuk analisis sentimen.

3.2.5. Tokenizing

Setelah *stopwords* dihapus dari data teks, selanjutnya data dilakukan *tokenizing*, yaitu mengubah teks menjadi token untuk mempermudah dalam proses selanjutnya yaitu *stemming*. Hasil dari proses *tokenizing* ditunjukkan pada tabel 19.

Tabel 19 memperlihatkan hasil dari proses tokenisasi, di mana setiap kata dalam teks diubah menjadi token.

Tabel 19. Hasil Proses *Tokenizing*

<i>stopword</i>	<i>tokenize</i>
still observer political expert	['still', 'observer', 'political', 'expert']
basic bajigur politician said wins win facing gentlemen back face majors questionnaire rights mk line	['basic', 'bajigur', 'politician', 'said', 'wins', 'win', 'facing', 'gentlemen', 'back', 'face', 'majors', 'questionnaire', 'rights', 'mk', 'line']
megamendung instead area house sirhuh please minthere residents need helped	['megamendung', 'instead', 'area', 'house', 'sirhuh', 'please', 'minthere', 'residents', 'need', 'helped']
gimmick	['gimmick']

3.2.6. Stemming

Langkah terakhir untuk *preprocessing* data adalah *stemming* dengan mengubah data teks menjadi kata dasar. Contoh hasil dari proses *stemming* ditunjukkan pada tabel 20.

Tabel 20. Hasil Proses *Stemming*

<i>tokenize</i>	<i>stemmed</i>
['still', 'observer', 'political', 'expert']	['still', 'observ', 'polit', 'expert']
['basic', 'bajigur', 'politician', 'said', 'wins', 'win', 'facing', 'gentlemen', 'back', 'face', 'majors', 'questionnaire', 'rights', 'mk', 'line']	['basic', 'bajigur', 'politician', 'said', 'win', 'face', 'gentelmen', 'back', 'face', 'major', 'questionnair', 'right', 'mk', 'line']
['megamendung', 'instead', 'area', 'house', 'sirhuh', 'please', 'minthere', 'residents', 'need', 'helped']	['megamendung', 'instead', 'area', 'hous', 'pleas', 'minther', 'resid', 'need', 'help']
['gimmick']	['gimmick']

Tabel 20 menunjukkan hasil dari proses *stemming*, di mana kata-kata yang telah diubah menjadi token kemudian dikembalikan ke bentuk kata dasar.

3.3. Analisis Sentimen

Untuk menganalisis sentimen data yang diperoleh sebelumnya. Dilakukan beberapa proses dan pelatihan model menggunakan data latih terlebih dahulu sehingga model terbiasa dengan bahasan politik.

3.3.1. Persiapan Data Latih

Pada tahap persiapan data latih, label teks, yang mewakili kategori atau sentimen, diubah menjadi representasi numerik menggunakan teknik *LabelEncoder*. Selanjutnya data dibagi menjadi dua kumpulan set yaitu set pelatihan dan set pengujian. Pembagian ini dilakukan dengan menggunakan fungsi *train_test_split* dari pustaka *scikit-learn*.

3.3.2. Hyperparameters

Parameter ditentukan untuk mengoptimalkan kinerja model yang digunakan dalam penelitian. *Hyperparameters* yang digunakan ditunjukkan pada tabel 21.

Tabel 21. Set *Hyperparameters*

Parameter	Jumlah
vocab_size	300
embedding_dim	100
max_length	200

Penjelasan dari parameter pada tabel 21 adalah sebagai berikut:

- *vocab_size* membatasi jumlah kata yang akan dipelajari dan diwakili oleh model. Parameter ini menentukan jumlah kata unik yang akan dipertimbangkan oleh model dalam proses *tokenizing*. *Tokenizing* adalah langkah awal

dalam pemrosesan bahasa alami di mana teks dipecah menjadi unit-unit yang lebih kecil (token) seperti kata-kata.

- *embedding_dim* menentukan dimensi vektor *embedding*. Vektor *embedding* adalah representasi numerik dari kata-kata yang dipelajari oleh model. Dimensi *embedding* menentukan seberapa kompleks dan bernuansa representasi kata-kata ini.
- *max_length* membatasi panjang kalimat yang dapat dipelajari dan diprediksi oleh model. Parameter ini menentukan panjang maksimum urutan teks yang akan diproses oleh model. Urutan teks adalah deretan kata yang membentuk sebuah kalimat atau paragraf.

3.3.3. Tokenizing dan Padding Data Latih

Tahap *tokenizing* dan *padding* dilakukan untuk memastikan semua kalimat pelatihan diubah menjadi *string*, membuat *tokenizer* dan *fit* pada kalimat pelatihan. Selain itu pada tahap ini dilakukan konversi kalimat pelatihan menjadi urutan angka dengan *sequences* dan juga dilakukan *padding* dengan menambahkan token khusus (biasanya token 0) ke urutan yang lebih pendek. Hal ini dilakukan untuk memastikan bahwa model dapat memproses semua kalimat dengan cara yang sama, terlepas dari panjangnya. Kemudian langkah yang sama diterapkan pada data pengujian.

3.3.4. Membuat Model LSTM

Pada pembuatan model LSTM dilakukan pendefinisian model dan juga penambahan lapisan-lapisan pada model. Lapisan model memuat definisi model, lapisan *Embedding*, lapisan LSTM *Bidirectional*, dan lapisan *Dense*. Penjelasan dari lapisan-lapisan tersebut adalah sebagai berikut.

- Definisi Model: Model didefinisikan menggunakan kelas *keras.Sequential* dari pustaka *Keras*. Kelas ini memungkinkan pembuatan model sekuensial, di mana lapisan disusun secara berurutan.
- Lapisan *Embedding*: Lapisan pertama adalah *Embedding* yang mengubah kata-kata dalam urutan *input* menjadi representasi vektor yang padat. Dimensi vektor *embedding* ditentukan oleh parameter *embedding_dim*, dan panjang maksimum urutan *input* ditentukan oleh *input_length*.
- Lapisan LSTM *Bidirectional*: Lapisan selanjutnya adalah *Bidirectional* yang membungkus lapisan LSTM. Lapisan LSTM adalah jenis jaringan saraf berulang (RNN) yang cocok untuk memproses data sekuensial seperti teks. Lapisan *Bidirectional* memungkinkan LSTM memproses urutan *input* di kedua arah, menangkap ketergantungan maju dan mundur dalam teks.

- Lapisan *Dense*: Model kemudian memiliki dua lapisan padat (*dense*):
 - Lapisan pertama memiliki 24 unit dan menggunakan fungsi aktivasi '*relu*'.
 - Lapisan kedua memiliki 1 unit dan menggunakan fungsi aktivasi '*sigmoid*'. Tujuan lapisan padat ini adalah untuk mengubah keluaran lapisan LSTM menjadi nilai prediksi tunggal. Fungsi aktivasi '*relu*' memperkenalkan non-linearitas ke model, memungkinkannya mempelajari hubungan yang lebih kompleks antara *input* dan *output*. Fungsi aktivasi '*sigmoid*' digunakan pada lapisan akhir untuk menghasilkan nilai probabilitas antara 0 dan 1, yang dapat diinterpretasikan sebagai probabilitas *input* termasuk dalam kelas tertentu.

3.3.5. Kompilasi dan Ringkasan Model

Kompilasi model adalah langkah di mana model pembelajaran mesin dikonfigurasi dengan beberapa parameter penting sebelum proses pelatihan dimulai. Dalam tahap ini, model diterjemahkan ke dalam bentuk yang dapat dipahami oleh mesin untuk dijalankan secara efektif pada perangkat keras yang tersedia. Sedangkan ringkasan model adalah representasi visual atau teks dari arsitektur model yang menunjukkan struktur model secara detail, termasuk informasi tentang jumlah *layer*, jenis *layer*, ukuran input dan output pada setiap *layer*, serta jumlah parameter yang dapat dilatih (*trainable parameters*) pada setiap *layer*. Hasil ringkasan model ditunjukkan pada tabel 22.

Tabel 22. Hasil ringkasan model

Layer (type)	Output Shape	Param #
<i>embedding_2</i> (<i>Embedding</i>)	(None, 150, 64)	960,000
<i>bidirectional_2</i> (<i>Bidirectional</i>)	(None, 128)	66,048
<i>dense_4</i> (<i>Dense</i>)	(None, 24)	3,096
<i>dense_5</i> (<i>Dense</i>)	(None, 1)	25

Tabel 22 menggambarkan arsitektur model jaringan saraf yang terdiri dari beberapa lapisan. Lapisan pertama adalah *embedding* yang menghasilkan *output* dengan dimensi 150 (panjang urutan *input*) dan 64 (dimensi *embedding*), memiliki 960.000 parameter. Selanjutnya, lapisan *bidirectional* LSTM menghasilkan *output* dengan dimensi 128 dan memiliki 66.048 parameter. Kemudian, lapisan *dense* pertama menghasilkan *output* dengan 24 unit dan memiliki 3.096 parameter. Akhirnya, lapisan *dense* terakhir menghasilkan *output* dengan satu unit, yang sering digunakan untuk tugas klasifikasi biner, dan memiliki 25 parameter. Secara keseluruhan, tabel ini menunjukkan struktur dan jumlah parameter dari model jaringan saraf yang digunakan.

3.3.6. Pelatihan Model

Langkah berikutnya adalah melatih model pada kumpulan data pelatihan. Model dilatih dengan jumlah *epoch* sebanyak 20 dengan *validation_split*

0,1 atau menggunakan data latih sebagai pengujian sebanyak 10%.

3.3.7. Implementasi Model

Setelah model dilatih dengan 20 kali *epoch*, langkah terakhir adalah menerapkan model untuk memprediksi sentimen pada teks hasil pemilu yang telah dikumpulkan. Hasil dari sentimen analisis ditunjukkan pada tabel 23.

Tabel 23. Hasil Analisis Sentimen

Text	Hasil Sentimen Analisis
still observ polit expert	Positive
basic bajigur politician said win win face	Positive
gentelmen back face major questionair right	Negative
mk line	Negative
megamendung instead area hous sirhuh pleas	Positive
minther resid need help	Positive
gimmick	Negative

Tabel 23 menampilkan hasil analisis sentimen yang diterapkan pada model, menghasilkan klasifikasi sentimen positif dan negatif dari data uji.

3.4. Analisis Korelasi

Setelah menganalisis sentimen publik terhadap hasil pemilu, langkah berikutnya adalah mengevaluasi korelasi antara sentimen yang dianalisis dan perubahan harga saham. Data sentimen dikelompokkan berdasarkan tanggal kemudian dihitung nilai rata-rata *prediction*-nya. Pada data saham juga dilakukan hal yang sama, dengan mengelompokkan berdasarkan tanggal dan dihitung rata-rata harga terakhir atau penutupan. Setelah itu, panjang data disamakan sehingga dapat dihitung korelasinya. Dari proses analisis yang telah dilakukan dihasilkan nilai korelasi *Spearman* sebesar 0,402 dan nilai p atau nilai signifikansi sebesar 0,046.

3.5. Evaluasi

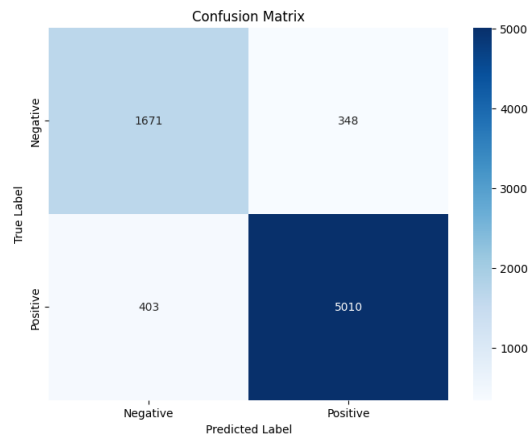
Tahap evaluasi dilakukan untuk menghitung angka akurasi, presisi, dan *recall* pada analisis sentimen dan mengevaluasi koefisien korelasi serta nilai p pada analisis korelasi.

3.5.1. Evaluasi Analisis Sentimen

Evaluasi hasil analisis sentimen dilakukan dengan menggunakan *confusion matrix*. Hasil *confusion matrix* dari analisis sentimen yang dilakukan ditunjukkan pada gambar 2.

Dari *confusion matrix* pada gambar 2, dapat dicari nilai akurasi, presisi, dan *recall* dengan menggunakan nilai pada matriks tersebut. Beberapa nilai tersebut adalah sebagai berikut:

- TP = 5010.
- FP = 348.
- TN = 1671.
- FN = 403.



Gambar 2. Confusion matrix

Dari nilai *confusion matrix* tersebut dapat dihitung nilai akurasi, presisi, dan *recall* sebagai berikut:

1. Akurasi

Berdasarkan persamaan 1, nilai akurasi dihitung sebagai berikut:

$$\begin{aligned} \text{Akurasi} &= (TP + TN) / (TP + FP + TN + FN) \\ \text{Akurasi} &= (5010 + 1671) / (5010 + 348 + 1671 + 403) \\ \text{Akurasi} &= 6681 / 7432 \\ \text{Akurasi} &= 0.900 \\ \text{Akurasi} &= 90\% \end{aligned}$$

Dengan demikian, nilai akurasi yang diperoleh adalah 90%.

2. Presisi

Berdasarkan persamaan 2, nilai presisi dihitung sebagai berikut:

$$\begin{aligned} \text{Presisi} &= TP / (TP + FP) \\ \text{Presisi} &= 5010 / (5010 + 348) \\ \text{Presisi} &= 5010 / 5358 \\ \text{Presisi} &= 0.936 \\ \text{Presisi} &= 93.6\% \end{aligned}$$

Dengan demikian, nilai presisi yang diperoleh adalah 93,6%.

3. Recall

Berdasarkan persamaan 3, nilai *recall* dihitung sebagai berikut:

$$\begin{aligned} \text{Recall} &= TP / (TP + FN) \\ \text{Recall} &= 5010 / (5010 + 403) \\ \text{Recall} &= 5010 / 5413 \\ \text{Recall} &= 0.927 \\ \text{Recall} &= 92.7\% \end{aligned}$$

Dengan demikian, nilai akurasi yang diperoleh adalah 92,7%.

Interpretasi perhitungan akurasi, presisi, dan *recall* adalah,

1. Akurasi: Nilai akurasi 0,900 menunjukkan bahwa model klasifikasi memprediksi dengan benar sekitar 90% dari semua kasus.
2. Presisi: Nilai presisi 0,936 menunjukkan bahwa dari semua kasus yang diklasifikasikan sebagai positif oleh model, 93,6% benar-benar positif.
3. Recall: Nilai *recall* 0,927 menunjukkan bahwa nilai *recall* mendapat nilai positif sebanyak 92,7% dan dari semua kasus yang sebenarnya

positif, model klasifikasi hanya melewati 7,3% kasus.

Berdasarkan interpretasi di atas, model klasifikasi ini memiliki akurasi, presisi, dan *recall* yang tinggi. Hal ini menunjukkan bahwa model klasifikasi LSTM sangat baik dalam mengidentifikasi sentimen positif dan negatif.

3.5.2. Evaluasi Analisis Korelasi

Evaluasi korelasi dinilai dari nilai koefisien korelasi dan nilai signifikansi untuk melihat apakah korelasi yang ada merupakan sebuah kebetulan atau bukan. Berikut interpretasi hasil nilai korelasi:

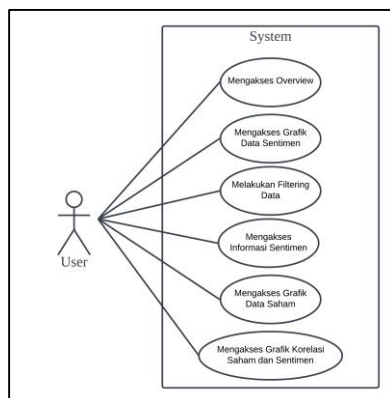
- Nilai koefisien korelasi *Spearman* sebesar 0,402 menunjukkan adanya hubungan negatif moderat antara dua variabel yang dianalisis. Korelasi positif berarti bahwa ketika satu variabel meningkat, variabel lainnya cenderung meningkat.
- Nilai *p* sebesar 0,046 adalah ukuran dari signifikansi statistik korelasi yang ditemukan. Ini menunjukkan bahwa ada kurang dari 4,6% kemungkinan bahwa korelasi yang diamati adalah hasil dari kebetulan. Hal ini berarti dapat dipastikan dengan cukup yakin bahwa korelasi positif yang diamati antara dua variabel bukanlah hasil kebetulan.

3.6. Dashboard

Untuk menampilkan data yang digunakan dalam penelitian ini serta untuk menentukan apakah data-data tersebut benar-benar berkorelasi, data disajikan dalam bentuk visualisasi yang digabungkan ke dalam *dashboard* interaktif. Dengan demikian, data akan lebih mudah dipahami bahkan oleh orang yang tidak terbiasa membaca data. Pada tahap pembuatan *dashboard* melalui beberapa tahapan sebagai berikut.

3.6.1. Use case

Use case dilakukan untuk mengidentifikasi dan mendefinisikan kebutuhan sistem berdasarkan interaksi antara pengguna (aktor) dan sistem tersebut. *Use case* pada sistem visualisasi penelitian ini ditunjukkan pada gambar 3.



Gambar 3. Use case Sistem Visualisasi

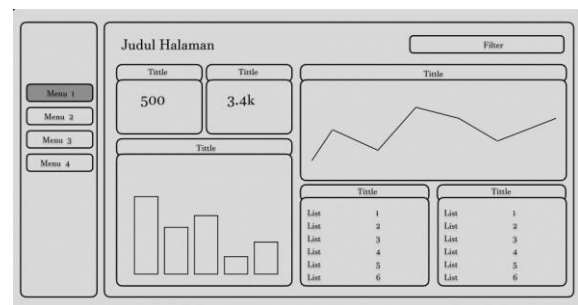
Diagram *use case* pada gambar 3 menunjukkan interaksi antara seorang pengguna (*User*) dan sistem untuk berbagai fungsi. *User* dapat mengakses *overview* untuk mendapatkan gambaran umum, melihat grafik data sentimen dan data saham untuk memahami tren, serta mengakses informasi sentimen untuk detail lebih mendalam. Selain itu, *user* juga dapat melakukan *Filtering* data untuk menyaring informasi sesuai kebutuhan dan melihat grafik korelasi antara saham dan sentimen untuk menganalisis hubungan antara dua data tersebut. Semua fungsi ini dirancang untuk memberikan pengguna alat yang diperlukan untuk analisis data yang lebih efektif dan mendalam.

3.6.2. Perancangan

Pada tahap perancangan, dibuat *mockup* atau rancangan tabel. Untuk pembagian visualisasi adalah data dengan jenis *time series* dibuat menjadi *line chart*, data dengan jenis kategori dibuat menjadi *bar chart*, sedangkan untuk perincian lebih lanjut bisa ditampilkan pada list.

3.6.2.1. Perancangan Halaman Overview

Perancangan halaman *overview* ditunjukkan pada gambar 4.

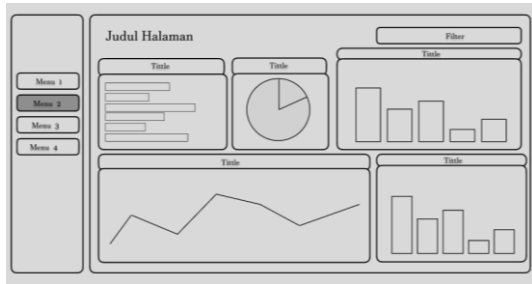


Gambar 4. Rancangan Halaman Overview

Gambar 4 menampilkan desain halaman *overview* yang terdiri dari beberapa elemen penting. Di sebelah kiri, terdapat empat tombol menu yang berfungsi sebagai navigasi antar halaman atau fitur dalam *dashboard*. Di bagian atas, terdapat judul halaman yang menunjukkan topik utama dari *dashboard*, serta sebuah kotak filter di pojok kanan atas untuk menyaring data sesuai dengan kriteria tertentu. Di tengah *dashboard*, ada beberapa kotak data yang menampilkan informasi penting berupa angka, serta dua jenis grafik, grafik batang di bagian bawah dan grafik garis di kanan atas yang menyajikan visualisasi data untuk menunjukkan tren atau perubahan dari waktu ke waktu. Di bagian kanan bawah, terdapat dua kolom yang berisi daftar item yang mungkin berisi poin-poin penting atau informasi terkait lainnya. Desain ini dirancang untuk menampilkan informasi secara ringkas dan visual, sehingga memudahkan pengguna dalam memahami data yang disajikan.

3.6.2.2. Perancangan Halaman Visualisasi Data

Perancangan halaman visualisasi data ditunjukkan pada gambar 5.

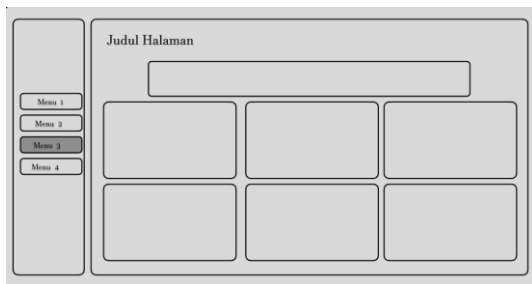


Gambar 5. Rancangan Halaman Visualisasi Data

Gambar 5 merupakan rancangan halaman visualisasi data pada *dashboard*. Terdapat beberapa elemen visualisasi data, termasuk diagram batang horizontal di kiri atas, diagram lingkaran di kanan atas, dan dua grafik batang serta grafik garis di bagian bawah. Masing-masing elemen ini menampilkan data dalam berbagai bentuk grafik untuk memudahkan analisis visual.

3.6.2.3. Perancangan Halaman Informasi

Perancangan halaman informasi data ditunjukkan pada gambar 6.

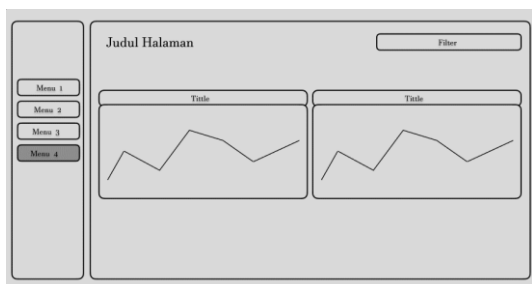


Gambar 6. Rancangan Halaman Informasi

Gambar 6 menampilkan rancangan halaman informasi pada *dashboard* berisikan beberapa kotak yang disusun dalam dua baris, yang kemungkinan digunakan untuk menampilkan berbagai informasi. Desain ini memberikan ruang yang luas untuk menampilkan berbagai konten secara terorganisir.

3.6.2.4. Perancangan Halaman Korelasi

Perancangan halaman visualisasi data ditunjukkan pada gambar 7.



Gambar 7. Rancangan Halaman Korelasi

Rancangan gambar 7 menampilkan hasil analisis korelasi. Pengguna dapat memilih variabel yang ingin dibandingkan melalui menu yang tersedia. Hasil korelasi akan ditampilkan dalam bentuk [misalnya, tabel, grafik]. Fitur filter memungkinkan pengguna untuk menyaring data berdasarkan kriteria tertentu.

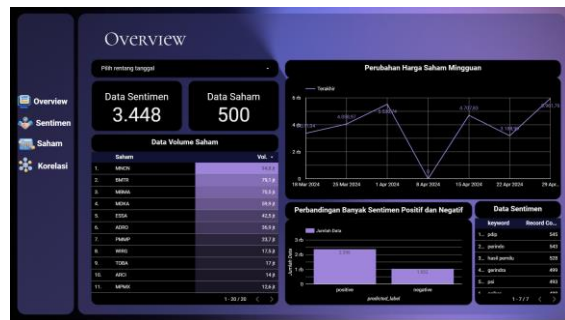
3.6.3. Implementasi

Pada tahap implementasi, rancangan yang telah dibuat kemudian direalisasikan pada *looker studio*. *Dashboard* visualisasi dibagi menjadi beberapa bagian yang memuat adanya data sentimen, data saham, dan korelasi keduanya. Untuk melihat *dashboard* visualisasi yang telah diimplementasikan dapat diakses melalui tautan <https://lookerstudio.google.com/s/vRN2x1Gi2Z8>.

Berikut adalah beberapa dokumentasi *dashboard* yang telah dibuat:

3.6.3.1. Overview

Bagian ini memuat visualisasi jumlah kedua data, perubahan harga saham, volume saham, perbandingan sentimen positif dan negatif, serta rincian jumlah sentimen untuk setiap kata kunci. Selain itu, terdapat filter untuk menyaring data berdasarkan rentang tanggal yang diinginkan. Visualisasi data pada halaman *overview* ditunjukkan pada gambar 8.



Gambar 8. Halaman Visualisasi Overview

Visualisasi yang disajikan pada gambar 14 memberikan gambaran menyeluruh mengenai pergerakan harga saham beberapa perusahaan dalam periode waktu tertentu. Selain itu, visualisasi ini juga mengukur sentimen pasar terhadap perusahaan-perusahaan tersebut dengan menganalisis kata kunci tertentu dalam berita atau diskusi *online*. Hubungan antara harga saham dan sentimen pasar ditampilkan secara visual, memungkinkan kita untuk melihat apakah terdapat korelasi positif atau negatif. Dengan kata lain, kita dapat melihat apakah peningkatan sentimen positif terhadap suatu perusahaan berdampak pada kenaikan harga sahamnya. Selain itu, visualisasi ini juga menyajikan data volume perdagangan, yang memberikan indikasi tentang aktivitas jual beli saham. Dengan menganalisis visualisasi ini, kita dapat memperoleh pemahaman

yang lebih baik mengenai faktor-faktor yang mempengaruhi pergerakan harga saham suatu perusahaan, serta tren pasar secara keseluruhan.

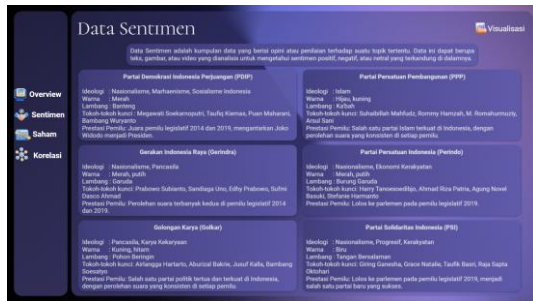
3.6.3.2. Data Sentimen

Visualisasi pada halaman data saham ditunjukkan pada gambar 9 dan untuk halaman informasi partai dapat ditunjukkan pada gambar 10.



Gambar 9. Halaman Visualisasi Data Sentimen

Pada gambar 9 memuat perincian mendalam mengenai data sentimen, seperti perbandingan antara sentimen positif dan negatif, perbandingan sentimen berdasarkan kata kunci, jumlah sentimen berdasarkan waktu, serta perbandingan jumlah data sentimen per kata kunci. Selain itu, terdapat filter untuk melihat data berdasarkan kata kunci tertentu dan rentang waktu tertentu.

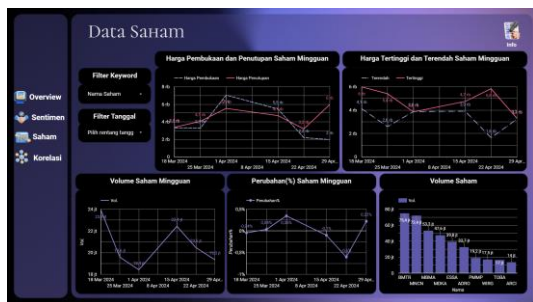


Gambar 10. Halaman Informasi Data Sentimen

Pada gambar 10 terdapat bagian info yang memberikan informasi seputar partai yang menjadi kata kunci pada penelitian ini.

3.6.3.3. Data Saham

Visualisasi pada halaman data saham dapat dilihat pada gambar 11.



Gambar 11. Halaman Visualisasi Data Saham

Visualisasi data saham pada gambar 11 memuat perincian lebih mendalam mengenai data saham, seperti pergerakan harga pembukaan, penutupan, tertinggi, dan terendah dari waktu ke waktu dalam bentuk mingguan, volume saham mingguan, perubahan (%) harga, serta volume saham untuk setiap nama saham. Selain itu, terdapat filter untuk melihat data saham berdasarkan nama saham dan rentang waktu tertentu.

3.6.3.4. Korelasi

Visualisasi pada halaman korelasi dapat dilihat pada gambar 12.



Gambar 12. Halaman Visualisasi Korelasi

Visualisasi korelasi pada gambar 12 memuat dua grafik visualisasi yang menunjukkan perubahan harga saham dari waktu ke waktu dan perbandingan jumlah sentimen positif dan negatif dari waktu ke waktu. Dengan demikian, audiens dapat melihat korelasi antara kedua data tersebut melalui grafik yang disajikan. Sebagai referensi hubungan saham dan kata kunci mana yang berkaitan disajikan tabel antara nama saham dan juga kata kunci yang terkaitnya. Selain itu, terdapat filter untuk menyaring data saham dan data sentimen berdasarkan kata kunci yang akan dipantau korelasinya, serta filter untuk melihat data dalam rentang tanggal tertentu.

4. DISKUSI

Penelitian ini berhasil menganalisis korelasi antara sentimen publik terhadap hasil Pemilu 2024 dengan pergerakan harga saham pelaku politik di Indonesia. Dari hasil yang diperoleh, beberapa poin penting dapat dibahas lebih lanjut.

Pertama, penggunaan model *Long Short-Term Memory* (LSTM) dalam analisis sentimen menunjukkan kinerja yang cukup baik dengan akurasi mencapai 90%, presisi sebesar 93,6%, dan *recall* sebesar 92,7%. Hasil ini mengindikasikan bahwa model LSTM mampu mengklasifikasikan sentimen publik dengan tingkat ketepatan yang tinggi. Namun, meskipun model ini cukup andal, ada kemungkinan bahwa model ini masih bisa ditingkatkan, terutama dalam menangkap nuansa sentimen yang lebih kompleks atau ambigu dalam data.

Kedua, hasil analisis korelasi *Spearman* menunjukkan adanya korelasi positif yang moderat antara sentimen publik terhadap hasil pemilu dan

perubahan harga saham pelaku politik, dengan koefisien sebesar 0,402 dan nilai p sebesar 0,046. Temuan ini konsisten dengan teori bahwa sentimen publik dapat mempengaruhi pasar saham, terutama dalam konteks politik. Namun, korelasi yang moderat ini juga menunjukkan bahwa ada faktor lain yang mungkin berperan dalam perubahan harga saham, seperti kondisi ekonomi atau kebijakan pemerintah yang belum sepenuhnya terisolasi dalam penelitian ini.

Selanjutnya, pembuatan *dashboard* interaktif menggunakan Looker Data Studio berhasil memfasilitasi visualisasi dan interpretasi data yang lebih mudah bagi pengguna. *Dashboard* ini memungkinkan analisis data secara lebih dinamis, memberikan akses yang intuitif untuk melihat tren sentimen dan pergerakan harga saham secara bersamaan. Namun, tantangan yang dihadapi adalah memastikan *dashboard* tetap relevan dan up-to-date seiring dengan perkembangan data yang terus berubah.

Secara keseluruhan, penelitian ini memberikan wawasan yang berharga mengenai hubungan antara sentimen publik dan pasar saham dalam konteks politik Indonesia. Meskipun temuan ini signifikan, ada beberapa keterbatasan yang harus diakui. Misalnya, faktor-faktor eksternal yang mungkin mempengaruhi hasil belum sepenuhnya dikendalikan, dan data sentimen yang diperoleh mungkin memiliki bias karena keterbatasan dalam metode crawling data.

5. KESIMPULAN

Dari penelitian yang sudah dilaksanakan dengan judul “Analisis Korelasi Sentimen Hasil Pemilu 2024 dan Pergerakan Saham Pelaku Politik di Indonesia” menghasilkan beberapa temuan penting yang dapat disimpulkan sebagai berikut:

Penelitian ini menggunakan bahasa pemrograman Python 3.10 dengan bantuan library seperti Scikit-learn, Pandas, dan Numpy. Data sentimen publik terhadap hasil pemilu diperoleh melalui proses crawling dari Twitter menggunakan alat Harvest Tweet. Metode *Long Short-Term Memory* (LSTM) digunakan untuk analisis sentimen, yang kemudian dievaluasi dengan hasil akurasi sebesar 90%, presisi sebesar 93,6%, dan *recall* sebesar 92,7%. Model LSTM ini menunjukkan kemampuan yang baik dalam mengklasifikasikan sentimen publik terhadap hasil pemilu.

Untuk mengkorelasikan sentimen publik dengan perubahan harga saham pelaku politik di Indonesia, dilakukan analisis korelasi *Spearman*. Hasil analisis menunjukkan koefisien *Spearman* sebesar 0,402 dengan nilai p sebesar 0,046. Ini menandakan adanya korelasi positif yang moderat antara sentimen publik terhadap hasil pemilu dengan perubahan harga saham. Dengan kata lain, semakin positif sentimen publik, semakin besar kemungkinan

harga saham pelaku politik mengalami kenaikan, dan sebaliknya.

Untuk memudahkan interpretasi data, dibuat *dashboard* interaktif menggunakan Looker Data Studio. *Dashboard* ini menampilkan visualisasi data yang digunakan dalam penelitian, memungkinkan pengguna untuk mengakses dan memahami data secara lebih mudah dan intuitif.

DAFTAR PUSTAKA

- [1] M. Habibie *et al.*, “Analisis Gejolak Saham Pada Peristiwa Politik Pemilihan Presiden Amerika Serikat 2020 Sektor Food and Beverage,” vol. 27, no. 2, 2022.
- [2] Z. Aprilia, “Prabowo Menang Quick Count, IHSG Lompat-Asing Borong Rp2,73 T Saham RI,” CNBC Indonesia. Accessed: Feb. 26, 2024. [Online]. Available: <https://www.cnbcindonesia.com/market/20240216065123-17-514885/prabowo-menang-quick-count-ihsg-lompat-asing-borong-rp273-t-saham-ri>
- [3] APJII, “Pengguna Internet di Indonesia Meningkat di 2024,” <https://apjii.or.id/berita/d/pengguna-internet-di-indonesia-meningkat-di-2024>.
- [4] Y. Mao, Q. Liu, and Y. Zhang, “Sentiment analysis methods, applications, and challenges: A systematic literature review,” Apr. 01, 2024, *King Saud bin Abdulaziz University*. doi: 10.1016/j.jksuci.2024.102048.
- [5] A. A. Firdaus, A. Yudhana, I. Riadi, and Mahsun, “Indonesian presidential election sentiment: Dataset of response public before 2024,” *Data Brief*, vol. 52, Feb. 2024, doi: 10.1016/j.dib.2023.109993.
- [6] M. M. Abdelgwad, T. H. A. Soliman, A. I. Taloba, and M. F. Farghaly, “Arabic aspect based sentiment analysis using bidirectional GRU based models,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 9, pp. 6652–6662, Oct. 2022, doi: 10.1016/j.jksuci.2021.08.030.
- [7] A. Nurkholis and Styawati, “Analisis Perbandingan Algoritma LSTM dan Naive Bayes untuk Analisis Sentimen,” *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 8, no. 2, pp. 299–303, Nov. 2022.
- [8] Y. Romadhoni, K. Fahmi, and H. Holle, “Analisis Sentimen Terhadap PERMENDIKBUD No.30 pada Media Sosial Twitter Menggunakan Metode Naive Bayes dan LSTM,” vol. 7, no. 2, 2022.
- [9] A. Heinen and A. Valdesogo, “Spearman rank correlation of the bivariate Student t and scale mixtures of normal distributions,” *J Multivar Anal*, vol. 179, Sep. 2020, doi:

- 10.1016/j.jmva.2020.104650.
- [10] B. Wang, R. Wang, and Y. Wang, "Compatible matrices of *Spearman's* rank correlation," *Stat Probab Lett*, vol. 151, pp. 67–72, Aug. 2019, doi: 10.1016/j.spl.2019.03.015.
- [11] N. Das, B. Sadhukhan, R. Chatterjee, and S. Chakrabarti, "Integrating sentiment analysis with graph neural networks for enhanced stock prediction: A comprehensive survey," Mar. 01, 2024, *Elsevier Inc.* doi: 10.1016/j.dajour.2024.100417.
- [12] Gallan V, "LSTM (Long Short Term Memory)." Accessed: Mar. 07, 2024. [Online]. Available: <https://medium.com/bina-nusantara-it-division/lstm-long-short-term-memory-d29779e2ebf8>
- [13] R. C. Staudemeyer and E. R. Morris, "Understanding LSTM -- a tutorial into *Long Short-Term Memory* Recurrent Neural Networks," p. 19, Sep. 2019, [Online]. Available: <http://arxiv.org/abs/1909.09586>
- [14] W. Tanjung and R. Ramadani, "Analisis korelasi pendapatan masyarakat dengan politik uang dalam pemilu," *JRTI (Jurnal Riset Tindakan Indonesia)*, vol. 8, no. 1, p. 190, Sep. 2023, doi: 10.29210/30033202000.
- [15] W. Tanjung and R. Ramadani, "Analisis korelasi pendapatan masyarakat dengan politik uang dalam pemilu," *JRTI (Jurnal Riset Tindakan Indonesia)*, vol. 8, no. 1, p. 190, Sep. 2023, doi: 10.29210/30033202000.
- [16] M. M. Fahmy, "Confusion matrix in Binary Classification Problems: A Step-by-Step Tutorial," *ERJ*, 2022.
- [17] A. Vázquez-Ingelmo, F. J. García-Peñalvo, R. Therón, and M. Á. Conde, "Representing data visualization goals and tasks through meta-modeling to tailor information dashboards," *Applied Sciences (Switzerland)*, vol. 10, no. 7, Apr. 2020, doi: 10.3390/app10072306.
- [18] S. Song, Y. Wang, X. Wang, C. Lin, and K. Hu, "A deep learning-based approach to similarity calculation for UML use case models," *Expert Syst Appl*, vol. 251, Apr. 2024..