

## Analisis Sentimen Twitter terhadap Isu Penundaan Pemilu 2024 Menggunakan Metode Long Short-Term Memory (LSTM)

Andy Putra Prasetyo\*<sup>1</sup>, Christian Sri Kusuma<sup>1</sup>, Didih Rizki<sup>1</sup>

Universitas Muhammadiyah Malang

andiputra1315@gmail.com\*

### Abstrak

Media sosial adalah bentuk aktivitas sosial yang menggunakan jaringan online untuk berkomunikasi melalui teks, gambar, dan video. Salah satu platform media sosial yang tetap menjadi tempat ekspresi bebas bagi individu adalah Twitter. Penelitian ini mengkaji sentimen masyarakat di Twitter terkait isu penundaan pemilu 2024 di Indonesia. Fokus utama penelitian ini adalah untuk menganalisis pandangan publik terhadap isu tersebut menggunakan metode Long Short-Term Memory (LSTM) dalam analisis sentimen. Data yang digunakan berasal dari cuitan-cuitan yang terkait dengan isu penundaan pemilu di Twitter. Tujuan penelitian ini adalah untuk memahami sejauh mana sentimen masyarakat terhadap isu penundaan pemilu 2024, apakah bersifat positif, negatif, atau netral. Metode yang digunakan adalah LSTM, sebuah teknik deep learning yang digunakan untuk memproses data teks dalam urutan. Data yang digunakan mencakup 1758 sampel setelah melalui proses pra-pemrosesan. Hasil penelitian menunjukkan bahwa sebanyak 58,0% dari sampel menunjukkan sentimen positif, sementara 42,0% menunjukkan sentimen negatif. Model LSTM yang diterapkan berhasil mencapai tingkat akurasi sebesar 88% dalam mengklasifikasikan sentimen masyarakat di Twitter terkait isu penundaan pemilu. Kesimpulan penelitian ini adalah mayoritas masyarakat yang terlibat dalam percakapan di Twitter terkait isu penundaan pemilu 2024 menunjukkan sentimen positif. Namun, masih ada sebagian yang menunjukkan sentimen negatif. Tingkat akurasi model LSTM yang tinggi menunjukkan efektivitas metode ini dalam menganalisis data teks dari media sosial. Oleh karena itu, penelitian ini memberikan wawasan yang berharga tentang pandangan publik terhadap isu penting ini di era digital.

**Kata Kunci:** Long Short-Term Memory (LSTM), Media sosial, Pemilu 2024, Sentimen, Twitter

### Abstract

Social media is a form of social activity that utilizes online networks to communicate through text, images, and videos. One of the social media platforms that remains a space for individual expression is Twitter. This study examines public sentiment on Twitter regarding the postponement of the 2024 elections in Indonesia. The primary focus of this research is to analyze public opinion on this issue using the Long Short-Term Memory (LSTM) method in sentiment analysis. The data used is derived from tweets related to the postponement of the elections on Twitter. The aim of this study is to understand the extent of public sentiment towards the postponement of the 2024 elections, whether it is positive, negative, or neutral. The method employed is LSTM, a deep learning technique used to process text data in sequence. The data used includes 1758 samples after preprocessing. The research findings indicate that 58.0% of the samples exhibit positive sentiment, while 42.0% show negative sentiment. The LSTM model applied achieves an accuracy rate of 88% in classifying public sentiment on Twitter regarding the postponement of the elections. The conclusion of this study is that the majority of individuals engaged in conversations on Twitter regarding the postponement of the 2024 elections display positive sentiment. However, there is still a portion that expresses negative sentiment. The high accuracy rate of the LSTM model demonstrates the effectiveness of this method in analyzing text data from social media. Therefore, this research provides valuable insights into public perspectives on this important issue in the digital era.

**Keywords:** Long Short-Term Memory (LSTM), Social media, Election 2024, Sentiment, Twitter

---

## 1. Pendahuluan

Indonesia ialah negara yang mempunyai warga dengan populasi tinggi, dapat dikatakan besar bahkan beragam suku atau latar belakang. Untuk menjaga stabilitas dalam Negara Kesatuan Republik Indonesia (NKRI), diperlukan suatu pemerintahan yang bijak untuk dapat mewakili keragaman masyarakat atau komunitas, meliputi aspek geografis dan ideologi. Ada metode untuk menentukan pemimpin pemerintahan. Terdapat cara memilih pimpinan, baik di eksekutif maupun legislatif, yang mewakili suara masyarakat Indonesia adalah dengan melalui Pemilu atau biasa disebut Pemilihan Umum (1)

Selain itu, pemilihan umum juga berperan sebagai sarana untuk memperkuat tanggung jawab dan pengawasan rakyat atas pemerintah. Haywood menyatakan bahwa pemilu berjalan dalam dua dimensi dari rakyat menuju negara (bottom-up) dan sebaliknya, negara menuju rakyat (top-down). Untuk menjamin terselenggaranya pemilihan umum yang bersih dan adil, sangat penting adanya kerangka peraturan-undangan yang rinci mengenai prosedur pemilu serta memberikan perlindungan kepada penyelenggara, kandidat, pemilih, pengawas, dan seluruh masyarakat dari segala macam kejahatan seperti intimidasi, kekerasan, penyuaipan, pemalsuan, dan Berbagai keadaan lain yang dapat mengganggu integritas hasil pemilihan. Oleh karena itu, keadilan dan kejujuran dalam pemilu memerlukan aturan pemilu yang tegas dan penerapan hukum yang tegas dan efektif (2).

Pada era globalisasi saat ini, kemajuan pada bidang teknologi telah berkembang dengan sangat cepat. dan signifikan. Baik di perusahaan maupun di instansi pemerintahan, peran teknologi sangat penting. Salah satu peran penting dalam perkembangan teknologi adalah komputer, yang berperan sebagai perangkat yang berfungsi dengan efisien dan efektif dalam berbagai kegiatan. Saat ini, teknologi hadir dalam bentuk aplikasi yang mencoba untuk membantu pekerjaan manusia dengan baik dan mudah digunakan. Salah satu contohnya adalah aplikasi Twitter (3).

Twitter merupakan platform media sosial yang memungkinkan individu dari berbagai belahan dunia untuk menyampaikan pendapat mereka. Data yang dihasilkan di Twitter memiliki potensi besar untuk dianalisis karena dapat diekstraksi menjadi informasi yang berharga melalui teknik analisis opini (4). Menurut data dari We Are Social pada bulan Januari 2023, jumlah pengguna Twitter global mencapai 556 juta, mengalami peningkatan sebesar 27,4% dari tahun sebelumnya. Indonesia memiliki kontribusi sebanyak 24 juta pengguna Twitter. Pendapat mengenai berita, peluncuran produk, atau tren tertentu dapat ditemukan dengan mudah di Twitter.

Media sosial adalah bentuk aktivitas sosial yang menggunakan jaringan online untuk berkomunikasi melalui teks, gambar, dan video. Salah satu platform media sosial yang tetap menjadi tempat ekspresi bebas bagi individu adalah Twitter. Platform ini memudahkan pengguna untuk berdiskusi, berbisnis, dan memberikan komentar dengan bebas. Twitter digunakan oleh berbagai kalangan untuk memberikan respons yang netral, positif, atau negatif terhadap topik yang tengah tren (5).

Luhut mengklaim bahwa 110 juta suara warganet mendukung pemilu pada 11 Maret 2022. Tetapi, informasi ini menjadi viral esok harinya dan dilaporkan media arus utama, yang menghasilkan grafik pada 12 Maret. Puncak kedua dalam grafik, bernilai 100, terjadi pada 12 April 2022, yang dipicu perdebatan kolektif antara mahasiswa Universitas Indonesia dan Luhut mengenai klaim Big Data 110 juta yang mendukung penundaan pemilu. Kejadian ini sedang diberitakan oleh berbagai media di Indonesia (6). Isu tersebut memunculkan berbagai macam pendapat dari kalangan masyarakat di platform media sosial Twitter.

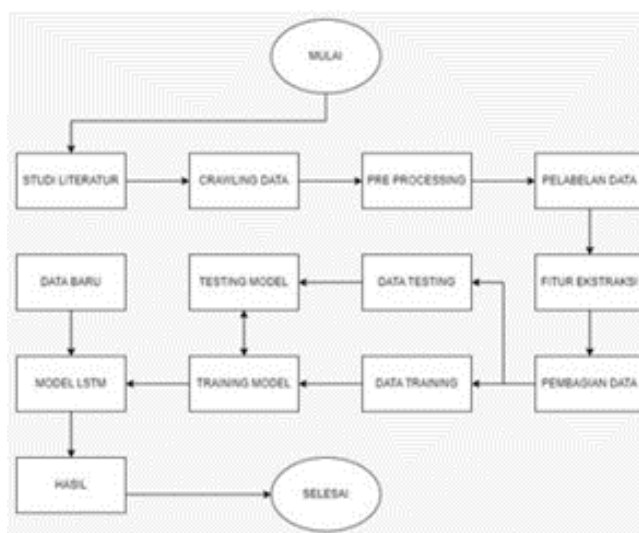
Menurut penelitian ini, direncanakan untuk melakukan analisis sentimen dengan memanfaatkan pendekatan Long Short-Term Memory (LSTM). Analisis sentimen ialah bentuk dari pendekatan (Natural Language Processing/NLP) yang tujuan untuk menelusuri sentimen atau suasana hati masyarakat terhadap produk atau topik tertentu. Juga dikenal sebagai penambangan opini, analisis sentimen melibatkan pengumpulan dan analisis pendapat yang terungkap dalam posting, blog, komentar, ulasan, dan cuitan. LSTM adalah turunan model dari Jaringan Saraf Rekuren (RNN) setelah dimodifikasi dengan menambahkan Unit memori yang dimana memori dapat menyimpan data atau informasi dalam kurun waktu yang lebih panjang (7). Penelitian ini bertujuan untuk memahami pandangan publik secara umum terkait dengan isu penundaan pemilu 2024.

## 2. Metode Penelitian

Studi literatur ialah pencarian referensi serta pengumpulan teori yang akan dipergunakan. Menelaah berbagai sumber seperti buku, literatur, catatan, jurnal, artikel, serta laporan mengenai penelitian LSTM (Long Short-Term Memory) guna melakukan penelitian untuk mengukur kepuasan pengguna (8).

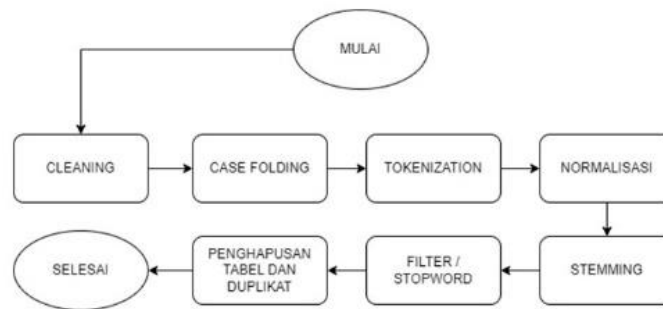
Crawling Data adalah pengambilan data yang dilakukan pada twitter. Pengumpulan data yang dilakukan menggunakan Application Programming Interface (API) Twitter dalam membuat sebuah sistem secara otomatis yang dapat mengambil data dari Twitter mempergunakan bahasa pemrograman Python. Data yang diambil termasuk informasi pengguna, tanggal pembuatan, dan konten yang dapat diekstraksi dari indeks yang tersedia di platform Twitter (9).

Crawling data pada twitter menggunakan sebuah commend line atau tools yaitu Tweet Harvest menggunakan Playwright untuk mengambil tweet dari pencarian Twitter dengan kata kunci dan rentang tanggal yang ditentukan. Tweet yang didapat akan disimpan dalam format file CSV. Untuk menggunakan Tweet Harvest, diperlukan perlu token otorisasi yang dapat diperoleh dengan login ke akun Twitter melalui browser dan mengekstrak authorization token, dan untuk library yang digunakan adalah library Tweepy (10). Library Tweepy sendiri adalah library Python yang digunakan untuk mengakses data witter, seperti tweet, profil pengguna, data tren, dan banyak lagi.



Gambar 1. Diagram Alur Penelitian

Pada Gambar 2 menjelaskan tentang tahapan pre-processing. Tahapan pertama yang dilakukan adalah tahapan cleaning yang dimana tahapan ini akan dibersihkan dari karakter yang tidak digunakan seperti angka, emoticon, dan tanda baca, seperti contoh kalimat berikut “Apakah km sdh makan ?” kalimat tersebut akan dirubah menjadi “Apakah km sdh makan”. Setelah itu data akan dibersihkan dari huruf kapital, jadi huruf kapital akan dirubah menjadi huruf kecil atau lowercase atau biasa di sebut Case Folding, maka kata tersebut akan menjadi “apakah km sdh makan”. Selanjutnya akan dilanjutkan ketahapan Tokenization yaitu melakukan pemotongan kalimat berdasarkan jumlah spasi yang ada dan akan dijadikan dalam bentuk kata sebagai contoh “apakah”, “km”, “sdh”, “makan”. Kemudian akan dilanjutkan ketahapan Normalization yang berarti mengubah kata singkatan atau kata tidak baku menjadi kata baku, yaitu kata “km” menjadi “kamu”, dan kata “sdh” menjadi “sudah”. Tahapan selanjutnya yaitu tahapan Stemming, tahapan ini akan merubah kata yang memiliki imbuhan kemudian akan menjadi kata dasar seperti “apakah” menjadi “apa”. Selanjutnya adalah tahapan Filtering atau stopward tahapan ini dilakukan untuk mengurangi kata yang tidak penting atau tidak menambahkan makna pada data seperti contoh “apa kam sudah makan” menjadi “sudah makan”. Kemudian tahapan Penghapusan table dan duplikat, tahapan ini menghapus duplikat data yang peneliti dapat. Hasil preprocessing ini digunakan untuk pembuatan model analisis.



Gambar 2. Diagram Pre-Processing

Pada tahapan labeling ini menggunakan kamus Lexicon dalam analisis sentimen dapat membantu secara otomatis menilai sentimen suatu teks yang melibatkan penggunaan kamus atau daftar kata-kata yang telah diberi label sentimen tertentu. Lexicon ini digunakan untuk menentukan sentimen suatu teks berdasarkan kata-kata yang terkandung di dalamnya (positif atau negative). Data sampel hasil labeling menggunakan Lexicon terlihat pada Tabel 1.

Tabel 1. Hasil Sentimen

Tweet	Sentimen
kecewa bapak anies cak imin sebagai calon wakil presiden nya karena cak imin adalah orang yang saya benci dia pernah suruh dan dukung tunda pemilu dengan dan telah pilih cak imin dari pada orang lain	0
apakah kalimat ini termasuk sentimen positif pemilu tetap berjalan sesuai jadwal tak ada tertunda pemilu	1

Pada langkah ini, setelah proses pelabelan data, data yang telah dilabelkan akan melalui proses fitur ekstraksi. Proses fitur ekstraksi ini bertujuan untuk mengubah data yang semula berupa kata-kata menjadi representasi data dalam bentuk vektor. Hal ini dilakukan agar data dapat dipahami dan diproses oleh sistem komputer.

Pada tahapan Pembagian Data ini dilakukan setelah selesai melakukan tahapan pre-processing data, data dari hasil preprocessing selanjutnya akan dilakukan ke tahapan pembagian data. Pembagian data tweet dilakukan dengan membagi dari tweet hasil pre-processing yang sudah melalui tahapan pelabelan, tahapan tersebut dibagi menjadi dua, yakni data train serta test. Data train digunakan melatih model klasifikasi yang dibuat, sedangkan data test digunakan menguji model klasifikasi yang di buat.

Tahapan training model LSTM ini adalah proses dimana model LSTM diajarkan untuk memahami pola dan informasi dalam data latihan yang diberikan agar proses prediksi bisa dilakukan dengan baik oleh model. Sebelum melakukan training model, dataset akan melalui proses pre-processing yang bertujuan merubah dataset menjadi bentuk vektor.

Pada tahapan Evaluasi Performa Model ini mempergunakan confusion matrix guna mengukur kinerja klasifikasi model. Confusion matrix dipergunakan sebagai perbandingan jumlah data yang diprediksi benar atau salah dengan jumlah data yang diprediksi sebenarnya. Hasil perbandingan yang didapat digunakan untuk menghitung akurasi, presisi recall, dan f1-score (11).

Analisis sentimen itu sendiri memiliki cakupan yang sangat luas. Sumber data yakni berita, ulasan, serta pesan singkat mengarah pada berbagai pendekatan pemodelan analisis sentimen yang berbeda. Pada tahapan ini data set didapatkan dari hasil crawling data dan telah melewati tahap pre-processing yang akan dilakukan tahap analisis sentiment dengan model yang telah dibangun untuk melihat sentiment terhadap topik "Analisis Sentimen Pada Twitter terhadap Isu Penundaan Pemilu 2024 Menggunakan Metode Long Short-Term Memory" selama penelitian ini dilakukan (8).

### 3. Hasil Penelitian dan Pembahasan

#### 3.1 Crawling Data

Pengambilan data mempergunakan google colabiratory atau biasa disebut google colab. Google colab merupakan cloud computing yang memungkinkan pengguna untuk menjalankan

kode Python secara langsung melalui web browser dengan fleksibilitas yang memperhatikan kebutuhan para programmer machine learning, analis big data, peneliti kecerdasan buatan, dan pembelajar Python. Untuk menggunakan Google Colaboratory dengank menggunakan web browser Chrome, Mozila Firefox, dan sebagainya, kemudian masuk ke platform google colab melalui link <https://colab.research.google.com/>.

Crawling data adalah proses mengumpulkan data dari platform media sosial, disini penulis mengumpulkan data menggunakan aplikasi Twitter untuk analisis sentiment, pemantauan tren, dan lainnya. Crawling data ini penulis menggunakan tweet harvest yang digunakan untuk proses pengumpulan data tentang penundaan pemilu, dengan berbagai kata kunci seperti "Isu\_pemilu\_2024", "Isu\_Penundaan\_Pemilu\_2024", "Pemilu\_2024", "Penundaan\_Pemilihan\_umum", "Penundaan\_Pemilu\_2024".



Gambar 3. Tampilan Crawling Data

Gambar 2 di atas adalah proses crawling data yang terkumpul sebanyak 2675 data yang dihasilkan dari berbagai macam kata kunci. Pada tahapan ini mendapatkan hasil data pada Tabel 2.

Tabel 2. Tabel Hasil Crawling

Waktu	User	Tweet
2023-11-26 19:15:54	FanduHolic	Mengapa jelang pemilu 2024 kok gada isu PKI. Karna yg menuduh dan yang dituduh skrng suatu dalam satu gerbong kereta yg sama.
2023-11-25 05:59:36	ubegebe1	Saya suka Pemilu 2024 ini. Ga kaya 2014 atau 2019 suasananya. Lbh santai. Ga ada isu kafiir, PKI dan apalah apalah.
2023-10-06 16:33:42	niki_fahri	Ada isu penundaan pemilu 2024 gara-gara masa jabatan Jokowi, katanya mau perpanjang. Heboh nih!". #DinastiPolitikJokowi #JumatBerkah #IndonesiaBukanMilikKeluarga"

### 3.2 Cleaning

Tujuan dari tahapan cleaning yang dimana tahapan ini akan membersihkan dari karakter atau elemen yang tidak digunakan bahkan dapat mengurangi arti dari sebuah kata seperti angka, emoticon, dan tanda baca (12).

Tabel 3. Tabel Hasil Cleaning

Sebelum	Sesudah
Mengapa jelang pemilu 2024 kok gada isu PKI. Karna yg menuduh dan yang dituduh skrng suatu dalam satu gerbong kereta yg sama.	Mengapa jelang pemilu kok gada isu PKI Karna yg menuduh dan yang dituduh skrng suatu dalam satu gerbong kereta yg sama
Saya suka Pemilu 2024 ini. Ga kaya 2014 atau 2019 suasananya. Lbh santai. Ga ada isu kafiir, PKI dan apalah apalah.	Saya suka Pemilu ini Ga kaya atau suasananya Lbh santai Ga ada isu kafiir PKI dan apalah apalah

Ada isu penundaan pemilu 2024 gara-gara masa jabatan Jokowi, katanya mau perpanjang. Heboh nih!". #DinastiPolitikJokowi #JumatBerkah #IndonesiaBukanMilikKeluarga"	Ada isu penundaan pemilu garagara masa jabatan Jokowi katanya mau perpanjang Heboh nih
--	--

Dalam tabel di atas bahwasanya seluruh unsur seperti titik, koma, URL, emotikon, angka, dan lainnya dihapus. Setelah tahap Cleaning data selesai, langkah selanjutnya adalah memasuki proses Case Folding.

### 3.3 Case folding

Pada proses Case Folding, dijelaskan bahwa huruf kapital dalam teks akan diubah menjadi huruf kecil atau lowercase. Hanya huruf 'a' hingga 'z' yang diterima dalam data Case Folding, sedangkan karakter selain huruf-huruf tersebut akan dihapus (13). Dapat melihat contoh proses Case Folding di Tabel 4.

Tabel 4. Tabel Hasil Case Folding

Cleaning	Case Folding
Mengapa jelang pemilu kok gada isu PKI Karna yg menuduh dan yang dituduh skrng suatu dalam satu gerbong kereta yg sama	mengapa jelang pemilu kok gada isu pki karna yg menuduh dan yang dituduh skrng suatu dalam satu gerbong kereta yg sama
Saya suka Pemilu ini Ga kaya atau suasananya Lbh santai Ga ada isu kafiir PKI dan apalah apalah	saya suka pemilu ini ga kaya atau suasananya lbh santai ga ada isu kafiir pki dan apalah apalah
Ada isu penundaan pemilu garagara masa jabatan Jokowi katanya mau perpanjang Heboh nih	ada isu penundaan pemilu garagara masa jabatan jokowi katanya mau perpanjang heboh nih

Pada tabel di atas bahwasanya tahapan Case Folding ialah mengubah huruf kapital menjadi huruf kecil 'a' sampai 'z'. Setelah tahapan Case Folding selanjutnya adalah Tokenization.

### 3.4 Tokenization

Tahapan Tokenization adalah tahapan dimana pemisahan kalimat menjadi pertoken kata (13) pada Tabel 5.

Tabel 5. Tabel Hasil Tokenization

Case folding	Tokenization
mengapa jelang pemilu kok gada isu pki karna yg menuduh dan yang dituduh skrng suatu dalam satu gerbong kereta yg sama	['mengapa', 'jelang', 'pemilu', 'kok', 'gada', 'isu', 'pki', 'karna', 'yg', 'menuduh', 'dan', 'yang', 'dituduh', 'skrng', 'suatu', 'dalam', 'satu', 'gerbong', 'kereta', 'yg', 'sama']
saya suka pemilu ini ga kaya atau suasananya lbh santai ga ada isu kafiir pki dan apalah apalah	['saya', 'suka', 'pemilu', 'ini', 'ga', 'kaya', 'atau', 'suasananya', 'lbh', 'santai', 'ga', 'ada', 'isu', 'kafiir', 'pki', 'dan', 'apalah', 'apalah']
ada isu penundaan pemilu garagara masa jabatan jokowi katanya mau perpanjang heboh nih	['ada', 'isu', 'penundaan', 'pemilu', 'garagara', 'masa', 'jabatan', 'jokowi', 'katanya', 'mau', 'perpanjang', 'heboh', 'nih']

Hasil dari Tokenization yang dimana kalimat tersebut akan dipisahkan menjadi pertoken kata. Setelah tahap Tokenization maka tahap selanjutnya adalah Normalisasi.

### 3.5 Normalisasi

Tahapan Normalisasi kata ini adalah tahapan merubah kata slang atau kata tidak baku menjadi kata baku seperti 'utk' menjadi 'untuk' (14). File normalisasi sendiri didapat dari GitHub yasirutomo dengan jumlah data normalisasi 1543, dan ditambahkan dengan beberapa kata normalisasi dari penulis yang tolat data menjadi 15 pada Tabel 6.

Tabel 6. Tabel Hasil Normalisasi

Tokenization	Normalisasi
[ 'mengapa', 'jelang', 'pemilu', 'kok', 'gada', 'isu', 'pki', 'karna', 'yg', 'menuduh', 'dan', 'yang', 'dituduh', 'skrng', 'suatu', 'dalam', 'satu', 'gerbong', 'kereta', 'yg', 'sama' ]	[ 'mengapa', 'jelang', 'pemilihan umum', 'kok', 'gada', 'isu', 'pki', 'karena', 'yang', 'menuduh', 'dan', 'yang', 'dituduh', 'sekarang', 'suatu', 'dalam', 'satu', 'gerbong', 'kereta', 'yang', 'sama' ]
[ 'saya', 'suka', 'pemilu', 'ini', 'ga', 'kaya', 'atau', 'suasananya', 'lbh', 'santai', 'ga', 'ada', 'isu', 'kafir', 'pki', 'dan', 'apalah', 'apalah' ]	[ 'saya', 'suka', 'pemilihan umum', 'ini', 'tidak', 'kaya', 'atau', 'suasananya', 'lebih', 'santai', 'tidak', 'ada', 'isu', 'kafir', 'pki', 'dan', 'apalah', 'apalah' ]
[ 'ada', 'isu', 'penundaan', 'pemilu', 'garagara', 'masa', 'jabatan', 'jokowi', 'katanya', 'mau', 'perpanjang', 'heboh', 'nih' ]	[ 'ada', 'isu', 'penundaan', 'pemilihan umum', 'karena', 'masa', 'jabatan', 'jokowi', 'katanya', 'mau', 'perpanjang', 'heboh', 'ini' ]

Dapat kita lihat pada tabel diatas bahwa proses Normalisasi kata ini merubah kata slang atau kata tidak baku menjadi kata baku. Setelah proses Normalisasi selanjutnya adalah tahapan Stamping.

### 3.6 Stamping

Tahap Stemming adalah tahap dimana kata dirubah menjadi kata dasar. Proses Stamping sendiri menggunakan library sastrawi. Pada tahap stamping ini pada Tabel 7.

Tabel 7. Tabel Hasil Stamping

Normalisasi	Stamping
[ 'mengapa', 'jelang', 'pemilihan umum', 'kok', 'gada', 'isu', 'pki', 'karena', 'yang', 'menuduh', 'dan', 'yang', 'dituduh', 'sekarang', 'suatu', 'dalam', 'satu', 'gerbong', 'kereta', 'yang', 'sama' ]	[ 'mengapa', 'jelang', 'pilih umum', 'kok', 'gada', 'isu', 'pki', 'karena', 'yang', 'tuduh', 'dan', 'yang', 'tuduh', 'sekarang', 'suatu', 'dalam', 'satu', 'gerbong', 'kereta', 'yang', 'sama' ]
[ 'saya', 'suka', 'pemilihan umum', 'ini', 'tidak', 'kaya', 'atau', 'suasananya', 'lebih', 'santai', 'tidak', 'ada', 'isu', 'kafir', 'pki', 'dan', 'apalah', 'apalah' ]	[ 'saya', 'suka', 'pilih umum', 'ini', 'tidak', 'kaya', 'atau', 'suasana', 'lebih', 'santai', 'tidak', 'ada', 'isu', 'kafir', 'pki', 'dan', 'apa', 'apa' ]
[ 'ada', 'isu', 'penundaan', 'pemilihan umum', 'karena', 'masa', 'jabatan', 'jokowi', 'katanya', 'mau', 'perpanjang', 'heboh', 'ini' ]	[ 'ada', 'isu', 'tunda', 'pilih umum', 'karena', 'masa', 'jabat', 'jokowi', 'kata', 'mau', 'panjang', 'heboh', 'ini' ]

Dapat kita lihat pada tabel diatas bahwa proses stemming ini merubah kata menjadi kata dasar seperti pada tabel diatas pada kata 'sukseskan' menjadi 'sukses' dan selanjutnya.

### 3.7 Stopword Removal atau Filtering

Tahapan Filtering atau biasa disebut Stopwords ini merupakan proses dimana kata yang dianggap tidak memiliki nilai atau kata yang tidak memiliki makna akan dihilangkan. Proses Filtering atau Stopword pada Tabel 8.

Tabel 8. Tabel Hasil Stopword Removal atau Filtering

Stamping	Stopword Removal atau Filtering
[ 'mengapa', 'jelang', 'pilih umum', 'kok', 'gada', 'isu', 'pki', 'karena', 'yang', 'tuduh', 'dan', 'yang', 'tuduh', 'sekarang', 'suatu', 'dalam', 'satu', 'gerbong', 'kereta', 'yang', 'sama' ]	[ 'jelang', 'pilih umum', 'gada', 'isu', 'pki', 'tuduh', 'tuduh', 'sekarang', 'suatu', 'satu', 'gerbong', 'kereta', 'sama' ]
[ 'saya', 'suka', 'pilih umum', 'ini', 'tidak', 'kaya', 'atau', 'suasana', 'lebih', 'santai', 'tidak', 'ada', 'isu', 'kafir', 'pki', 'dan', 'apa', 'apa' ]	[ 'suka', 'pilih umum', 'kaya', 'suasana', 'lebih', 'santai', 'isu', 'kafir', 'pki', 'apa', 'apa' ]
[ 'ada', 'isu', 'tunda', 'pilih umum', 'karena', 'masa', 'jabat', 'jokowi', 'kata', 'mau', 'panjang', 'heboh', 'ini' ]	[ 'isu', 'tunda', 'pilih umum', 'masa', 'jabat', 'jokowi', 'kata', 'mau', 'panjang', 'heboh' ]

Pada tabel diatas kata yang tidak bermakna akan dihilangkan seperti maka', 'akan', 'yang', 'untuk', 'dan', 'juga', 'dari', 'di', 'ya', serta 'kan'.

**3.8 Penghapusan Data Duplikat dan Data Kosong**

Tahapan ini adalah tahap dimana data yang sama dan data kosong akan dihapus dari data set, hal ini berfungsi untuk menghindari data yang sama atau data berulang dan data kosong.



Gambar 4. Penghapusan Data dan Duplikat Dan Data Kosong

Gambar 4 di atas merupakan gambar sebelum dan sesudah penghapusan duplikat dan data kosong, dari data awal sebanyak 2675 data kemudian menjadi 1758 data.

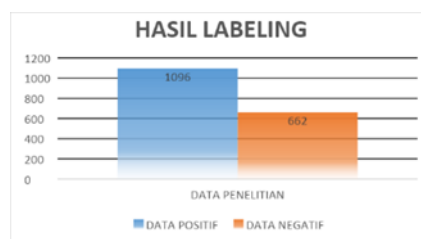
**3.9 Labeling**

Tahap Labeling ini peneliti melakukan pelabelan otomatis dengan menggunakan lexicon yang di dapat dari Github 'fajri91'. Proses labeling ini menghitung sentiment pada setiap kata dengan Polarity Score sebagai nilai dari setiap tweet dan polarity sebagai nilai yang didapat dari hasil labeling.

Tabel 9. Tabel Labeling

Filtering	Polatiry Score	Polarity
['jelang', 'pilih umum', 'gada', 'isu', 'pki', 'tuduh', 'tuduh', 'sekarang', 'suatu', 'satu', 'gerbong', 'kereta', 'sama']	3	1
['suka', 'pilih umum', 'kaya', 'suasana', 'lebih', 'santai', 'isu', 'kafir', 'pki', 'apa', 'apa']	2	1
['isu', 'tunda', 'pilih umum', 'masa', 'jabat', 'jokowi', 'kata', 'mau', 'panjang', 'heboh']	-1	0

Tabel diatas adalah tabel Labeling yang dibedakan menjadi 2, yaitu polarity yang bernilai 1 maka dinyatakan Positif, dan polarity 0 maka dinyatakan Negatif. Hasil labeling dapat kita lihat pada Gambar 5 dibawah.



Gambar 5. Hasil Labeling

Pada Gambar 5 diatas dataset yang telah melalui proses labeling mendapatkan data yang terbagi menjadi 2 data, yaitu positi berwarna biru sebanyak 1758 data, data negatif berwarna oranye sebanyak 662 data, dan data duplikat berwarna abu-abu sebanyak 1096 data yang dihapus.



### 3.10 Ekstraksi Fitur

Setelah dataset selesai melalui proses pre-processing tahap selanjutnya yaitu melakukan fitur ekstraksi. Tujuan dari fitur ekstraksi sebelum memasukkan data ke dalam model ialah mengubah data mentah ke dalam bentuk yang dapat dimengerti oleh system computer untuk membada dan mengolah kata agar dapat diproses oleh model atau merubah data yang sebelumnya berbentuk kata menjadi data berbentuk vektor. Pada klasifikasi data ini mempergunakan model LSTM dengan menggunakan fitur ekstraksi word embedding. Hasil dari menggunakan pendekatan atau fitur ekstraksi word embedding pada Gambar 6.

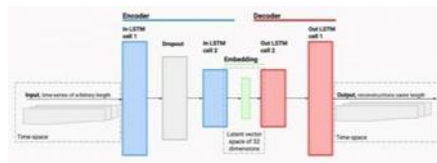
```
array([[ 0,  0,  0, ..., 663, 39, 16],
       [ 0,  0,  0, ...,  1,  2, 29],
       [ 0,  0,  0, ..., 208, 501, 85],
       ...,
       [ 0,  0,  0, ..., 48, 16, 215],
       [ 0,  0,  0, ..., 27,  1,  2],
       [ 0,  0,  0, ...,  4,  1,  2]], dtype=int32)
```

Gambar 6. Fitur Ekstraksi Word Embedding

Gambar di atas menunjukkan output dari penerapan word embedding yang mengubah kumpulan data tweet menjadi representasi vector.

### 3.11 Training Model LSTM

Pada training model data inputan digunakan untuk model klasifikasi yang dilakukan konversi dari data berbentuk kata menjadi data berbentuk vektor. Terdapat beberapa tahap dalam melakukan metode LSTM yang dilihat pada Gambar 7.



Gambar 7. Tahapan LSTM

Pada gambar training model LSTM diatas ada beberapa komponen dan tahapan pendukung untuk dilakukannya perhitungan: **Input:** Langkah pertama dalam arsitektur model, di mana data input dimasukkan ke dalam jaringan. Data input ini dapat berupa urutan kata atau token yang akan diproses oleh model LSTM.

**LSTM Cell 1:** Langkah kedua, di mana dilakukan pemrosesan data oleh LSTM cell pertama. LSTM cell ini bertanggung jawab untuk menangkap informasi jangka panjang dari urutan input yang diberikan.

**Dropout:** Dropout layer, yang ditempatkan setelah LSTM cell pertama. Dropout dipergunakan meminimalisir overfitting secara acak mematikan sebagian unit atau bobot selama proses pelatihan.

**LSTM Cell 2:** Langkah berikutnya, di mana dilakukan pemrosesan tambahan oleh LSTM cell kedua. LSTM cell kedua ini dapat menangkap informasi tambahan dari urutan input atau mungkin menghasilkan representasi yang lebih abstrak dari data.

**Embedding:** Langkah di mana representasi kata atau token diubah menjadi vektor padat dengan menggunakan embedding layer. Embedding membantu model dalam memahami makna relatif antara kata-kata dalam urutan input.

**Output LSTM 1:** Output dari LSTM cell pertama, yang mungkin merupakan representasi terstruktur dari bagian pertama dari urutan input.

**Output LSTM Cell 2:** Output dari LSTM cell kedua, yang mungkin merupakan representasi terstruktur dari seluruh urutan input setelah pemrosesan oleh kedua LSTM cell.

**Output:** Output akhir dari model, yang mungkin berupa prediksi kelas, nilai numerik, atau urutan kata-kata terjemahan, tergantung pada tugas spesifik dari model tersebut.

### 3.12 Evaluasi Peforma Model

Dalam tahap Evaluasi Performa Model, data dilatih menggunakan data training untuk mengevaluasi kinerja model yang telah dibuat seperti yang terlihat pada Gambar 8.



Gambar 8. Confusion Matriks

Pada Gambar 8 diperlihatkan contoh matriks dari pembagian data dengan rasio 9:1 yang menunjukkan hasil pengujian model dengan hasil data sebagai berikut:

1. True Positif (TP) yaitu data yang berlabel positif terprediksi sebagai data positif sebesar 66 data.
2. True Negatif (TN) yaitu data yang berlabel negatif terprediksi sebagai data negatif sebesar 60 data.
3. False Positif (FP) yaitu data yang berlabel positif terprediksi sebagai data negatif sebesar 11 data.
4. False Negatif (FN) yaitu data yang berlabel negatif terprediksi sebagai data positif sebesar 6 data.

Confusion Matrix digunakan sebagai indikator tingkat akurasi model yang telah dibuat dengan mempertimbangkan nilai dari Accuracy, Precision, Recall, serta F1-Score. Accuracy bertujuan untuk mengukur sejauh mana model dapat memprediksi dengan benar kelas atau label dari seluruh data yang ada. Akurasi dinyatakan sebagai presentase prediksi yang benar dari keseluruhan prediksi yang dilakukan model.

Selanjutnya dilakukan perhitungan Precision (Presisi): Juga dikenal sebagai Positive Predictive Value (PPV), mengukur seberapa baik model dalam memprediksi kelas positif secara tepat. Precision memberikan informasi tentang seberapa sering prediksi positif yang dibuat oleh model adalah benar.

Lalu dilakukan perhitungan Recall Juga dikenal sebagai Sensitivity atau True Positive Rate (TPR), mengukur seberapa baik model dalam mengidentifikasi semua instance yang benar-benar positif. Recall memberikan informasi tentang kemampuan model untuk menemukan semua sampel positif yang sebenarnya di dataset.

Kemudian perhitungan yang terakhir yaitu F1-score: yakni rata-rata harmonis antara recall serta precision. F1-score memberikan keseimbangan antara recall dan precision. Tujuan dari F1-score adalah untuk mempertimbangkan keduanya secara bersamaan, sehingga jika salah satunya rendah, F1-score juga akan rendah. Nilai dari evaluasi performa model dapat dilihat pada Tabel 10.

Tabel 10. Hasil Evaluasi Performa Model

Accuracy	Precision		Recall		F1-Score	
	0	1	0	1	0	1
0.88	0.91	0.86	0.85	0.92	0.88	0.89

Dari tabel diatas kita dapat menyimpulkan bahwa hasil dari Accuracy sebesar 88%, Precision positif 86%, Precision negatif 91%, Recall positif 92%, Recall negatif 85%, F1-Score positif 89% dan F1-Score negatif 88%.

### 3.13 Analisis Sentimen

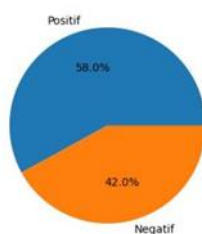
Analisis sentiment ini adalah tahapan dimana dataset diklasifikasikan kedalam sentimen positif dan sentimen negatif.

	Maktu	Tweets	Sentimen
0	2023-11-28 19:34:59	mana kampanye mulai calon siap bagi strategi	0
1	2023-11-28 19:24:20	isu wahl perempuan politik Indonesia kian san...	0
2	2023-11-28 09:54:14	bhayangkara bisa aman dan tertib masyarakat po...	0
3	2023-11-28 09:53:02	bhayangkara bisa aman dan tertib masyarakat po...	0
4	2023-11-28 08:59:47	hal sahabat live obrol newstroom hari bahas isu	1
...	...	...	...
1753	2024-02-16 07:35:00	panggil titik soeharto Prabowo Subianto jadi	1
1754	2024-02-16 07:35:00	wakapolda dy brigjen mantap adl vhid rumah p...	1
1755	2024-02-16 07:34:00	rumah sakit umum daerah roud moh anwar sumenep...	0
1756	2024-02-16 07:34:00	upaya proses pilh umum lancar transparan duka...	1
1757	2024-02-16 07:33:00	pilh umum bagi pilh umum buruk	0

1758 rows x 3 columns

Gambar 9. Hasil Preprocessing Data Baru

Pada Gambar 9 di atas terdapat 1758 data yang sudah melalui seluruh proses pre-processing kemudian melewati tahapan Long Short-Term Memory (LSTM). Hasil dari klasifikasi analisis sentimen ditunjukkan pada Gambar 10.



Gambar 10. Hasil Klasifikasi Analisis Sentimen

Pada gambar diatas mendapatkan hasil klasifikasi dari model LSTM ini dengan sentiment yang berlabelkan negative sebesar 42% dan sentiment berlabel positif sebesar 58%. Hasil dari Analisis Sentimen Pada Twitter terhadap Isu Penundaan Pemilu 2024 Menggunakan Metode Long Short-Term Memory (LSTM) ini menunjukkan bahwa lebih banyak sentiment yang mengarah pada sentimen positif.

## 4. Kesimpulan

Dalam hasil penelitian yang berjudul "Analisis Sentimen Pada Twitter terhadap Isu Penundaan Pemilu 2024 Menggunakan Metode Long Short-Term Memory (LSTM)" melibatkan 1758 sampel setelah proses pra-pemrosesan, ditemukan bahwa 58,0% menunjukkan sentimen positif, sementara 42,0% menunjukkan sentimen negatif. Model LSTM yang diterapkan berhasil mencapai tingkat akurasi sebesar 88%.

## Referensi

- [1] Rantau MI. Penguatan Sistem Presidensial Di Indonesia: Analisis Terhadap Undang Undang No 7 Tahun 2017 Tentang Pemilihan Umum. Jurnal Penelitian dan Karya Ilmiah. 2019;2.
- [2] Ramadhan MN. Evaluasi Penegakan Hukum Pidana Pemilu Dalam Penyelenggaraan Pemilu 2019. Vol. 2, Jurnal Adhyasta Pemilu. 2019.
- [3] Siregar LY, Nasution MIP. Perkembangan Teknologi Informasi Terhadap Peningkatan Bisnis Online. 2020;2(1):71–5. Available from: <https://doi.org/10.30606/hjimbhttp://journal.upp.ac.id/index.php/Hirarki>
- [4] Zuhdi AM, Utami E, Raharjo S. Analisis Sentiment Twitter Terhadap CAPRES Indonesia 2019 Dengan Metode K-NN. Jurnal INFORMA Politeknik Indonusa Surakarta. 2019;5:2442–7942.
- [5] Alita D, Fernando Y, Sulistiani H. Implementasi Algoritma Multiclass SVM Pada Opini Publik Berbahasa Indonesia Di Twitter. Jurnal TEKNOKOMPAK. 2020;14(2):86.
- [6] Alvin S. Analisis Framing Isu Penundaan Pemilu 2024 Di CNN Indonesia.Com Dan Kompas.Com. Jurnal SEMIOTIKA [Internet]. 2022;16(2):133–48. Available from: <http://journal.ubm.ac.id/>

- [7] Yahyadi A, Latifah F. Analisis Sentimen Twitter Terhadap Kebijakan Ppkm Di Tengah Pandemi Covid-19 Menggunakan Mode LSTM. *Journal of Information System, Applied, Management, Accounting and Research* [Internet]. 2022;6(2):464–70. Available from: <http://journal.stmikjayakarta.ac.id/index.php/jisamar>,
- [8] Amrustian MA, Widayat W, Wirawan AM. Analisis Sentimen Evaluasi Terhadap Pengajaran Dosen di Perguruan Tinggi Menggunakan Metode LSTM. *Jurnal Media Informatika Budidarma*. 2022 Jan 25;6(1):535–41.
- [9] Eka Sembodo J, Budi Setiawan E, Abdurahman Baizal Z. Data Crawling Otomatis pada Twitter. In *School of Computing, Telkom University*; 2016. p. 11–6.
- [10] Naufal MF, Kusuma SF. Analisis Sentimen pada Media Sosial Twitter Terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat Berbasis Deep Learning. *Jurnal Edukasi dan Penelitian Informatika*. 2022;8(1).
- [11] Normawati D, Prayogi SA. Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. Vol. 5, *Jurnal Sains Komputer & Informatika (J-SAKTI)*. 2021.
- [12] Kurniawan AA, Mustikasari M. Implementasi Deep Learning Menggunakan Metode CNN dan LSTM untuk Menentukan Berita Palsu dalam Bahasa Indonesia. 2020;5(4):2622–4615. Available from: <http://openjournal.unpam.ac.id/index.php/informatika544>
- [13] Talita AS, Wiguna A. Implementasi Algoritma Long Short-Term Memory (LSTM) Untuk Mendeteksi Ujaran Kebencian (Hate Speech) Pada Kasus Pilpres 2019. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*. 2019 Nov 2;19(1):37–44.
- [14] Cahyo PW, Aesy US. Perbandingan LSTM dengan Support Vector Machine dan Multinomial Na ve Bayes pada Klasifikasi Kategori Hoax. *Jurnal Transformatika*. 2023 Jan 20;20(2):23.