

Klasifikasi Opini Publik di Twitter Terhadap Bakal Calon Presiden Indonesia Tahun 2024 Menggunakan LSTM Secara *Realtime* Berbasis *Website*

Muhammad Rizki¹, Muhammad Fikri Hidayattullah², Dwi Intan Af'idah^{3*}

^{1,2,3}Program Studi Teknik Informatika, Politeknik Harapan Bersama

^{1,2,3}Jl. Mataram No.9 Pesurungan Lor, Kota Tegal, 52147, Indonesia

E-mail: iki11111998@gmail.com¹, fikri@poltektegal.ac.id², dwiintanafidah@poltektegal.ac.id³

Info Naskah:

Naskah masuk: 16 Mei 2023

Direvisi: 28 Juni 2023

Diterima: 29 Juni 2023

Abstrak

Banyaknya opini publik warganet Indonesia terkait bakal calon presiden Indonesia 2024 di media sosial Twitter sulit dianalisis. Penentuan klasifikasi oleh manusia membutuhkan tenaga ahli dan waktu lama. Oleh karena itu, diperlukan sistem visualisasi klasifikasi opini publik secara real-time. Penelitian sebelumnya hanya berakhir di evaluasi model, sedangkan penelitian ini berusaha menerapkan model terbaik pada website. Tujuan penelitian ini adalah menciptakan sistem memantau klasifikasi opini publik Twitter terhadap bakal calon presiden Indonesia 2024. Proses training menggunakan metode LSTM, menghasilkan model dengan akurasi 76%. Parameter yang diuji antara lain *batch size*, *dropout*, dan *learning rate*. Data berasal dari *crawling* Twitter dengan kata kunci Ganjar Pranowo, Anies Baswedan, dan Prabowo Subianto. Model LSTM diimplementasikan dalam website yang menampilkan dashboard dengan fitur peta warna tingkat sentimen positif bakal calon di setiap provinsi, jumlah klasifikasi sentimen, dan filter berdasarkan provinsi dan waktu.

Keywords:

public opinion;
presidential election;
LSTM;
real-time;
dashboard;

Abstract

The analysis of public opinions from Indonesian netizens regarding the potential presidential candidates for Indonesia in 2024 on Twitter is challenging. Human-based classification of the candidates on Twitter has limitations as it requires expertise and a considerable amount of time to process the data. Therefore, a system that provides real-time visualization of public opinion classification is necessary. Previous research only focused on model evaluation, while this study aims to implement the best model on a website. The objective of this research is to develop a system for monitoring the Twitter-based public opinion classification of the potential presidential candidates for Indonesia in 2024 within specific time frames. The training process utilizes the LSTM method, resulting in a model with an accuracy of 76%. Parameters such as batch size, dropout, and learning rate were tested. The data used in this study was obtained by crawling Twitter using the keywords Ganjar Pranowo, Anies Baswedan, and Prabowo Subianto. The LSTM model was then implemented in a website-based system that generates a dashboard with features such as a color-coded map displaying the highest levels of positive sentiment for each candidate in each province, the overall classification count for each candidate, and filters for sentiment classification based on province and specific time frames.

*Penulis korespondensi:

Dwi Intan Af'idah

E-mail: dwiintanafidah@poltektegal.ac.id

1. Pendahuluan

Presiden adalah orang yang memegang jabatan kepala negara dan kepala pemerintahan serta bertindak sebagai pemimpin suatu negara. Presiden Indonesia dipilih oleh rakyat dalam pemilihan presiden (pilpres) yang demokratis dan dilaksanakan setiap lima tahun sekali. Menjadi Presiden memiliki persyaratan yaitu tidak bisa menjadi Presiden jika sebelumnya telah menjabat dua periode berturut-turut sebagai Presiden. Jika demikian, Presiden Indonesia saat ini tidak akan dapat dipilih kembali sebagai Presiden pada pemilihan presiden berikutnya pada tahun 2024. [1]. Berdasarkan hal tersebut, muncul beberapa survei elektabilitas terhadap beberapa tokoh masyarakat dengan kualifikasi yang baik, sehingga memungkinkan menjadi calon presiden Indonesia pada Pilpres 2024, yang dapat ditemukan di baris berita, keempat lembaga survei tersebut adalah Lembaga survei Political Weather Station (PWS) [2], Indonesia Political Opinion (IPO) [3], Saiful Mujani Research and Consulting (SMRC) [4], Charta Politika [5]. Hasil keempat survei tersebut mendapatkan tiga tokoh publik dengan nilai elektabilitas terbesar yaitu Ganjar Pranowo, Anies Baswedan, dan Prabowo Subianto. Dimungkinkan untuk menyampaikan opini publik tentang hasil survei dan juga baris berita yang menerbitkan berita di jejaring sosial.

Jumlah pengguna media sosial di Indonesia saat ini mencapai 191,4 juta pengguna pada awal tahun 2022, dan jumlah pengguna media sosial di Indonesia diperkirakan akan tumbuh sebanyak 21 juta (+12,6 persen) pada tahun 2021-2022. Media sosial paling populer adalah Twitter dengan total 18,45 juta pengguna. Pengguna di Indonesia pada awal tahun 2022 [6]. Twitter juga digunakan oleh masyarakat untuk menyampaikan pendapat atau dengan nama lain Opini Publik. Opini yang disampaikan juga beragam tentang tokoh politik, sosial, olahraga atau peristiwa yang sedang viral dengan cara yang positif, negatif, dan netral. Klasifikasi berarti pengelompokan objek-objek yang memiliki sifat atau karakteristik yang sama ke dalam beberapa kelas [7]. Positif, negatif, dan netral merupakan Klasifikasi dari Opini Publik. Klasifikasi Opini Publik dari ke 3 tokoh yang menjadi bakal calon presiden Indonesia menjadi penting terhadap bakal calon presiden Indonesia tahun 2024 akan menghasilkan berbagai macam reaksi dan komentar dari masyarakat sehingga sulit untuk menganalisis tweet yang ada untuk bisa diolah menjadi informasi yang lebih baik. Penentuan Klasifikasi Opini Publik bakal calon presiden Indonesia tahun 2024 di twitter dengan bantuan manusia memiliki kelemahan karena membutuhkan tenaga ahli dan membutuhkan waktu lama untuk mengolah datanya, maka dari itu diperlukan website yang menampilkan *dashboard* yang dapat memberikan visualisasi klasifikasi opini publik bakal calon presiden Indonesia tahun 2024 pada twitter. Padahal, kebutuhan analisis opini publik dibutuhkan baik masyarakat luas ataupun tim sukses bakal calon [8].

Masalah lain muncul dalam kondisi riil sekarang masyarakat tidak bisa melihat opini publik yang positif atau negatif, sehingga masyarakat atau tim sukses tidak bisa melakukan *monitoring* perkembangan opini publik. Oleh karena itu dibutuhkan *website* yang menampilkan

dashboard yang dapat melakukan *monitoring* opini publik terhadap bakal calon presiden 2024. Klasifikasi Opini Publik atau yang biasanya dikenal klasifikasi sentimen merupakan bidang studi yang menganalisis pendapat, perasaan, penilaian, sikap, dan perasaan orang tentang bahasa tulis [9]. Sebagian besar kesulitan dalam analisis sentimen adalah karena jumlah data yang terlalu banyak untuk diproses secara manual. *Deep Learning* dapat mengatasi masalah dengan data yang jumlahnya besar dengan mudah, dibantu dengan model *deep learning* yang populer telah membuat masalah NLP jauh lebih mudah [10].

Pada beberapa penelitian [11], [12], [13], algoritma *machine learning* SVM mendapatkan performa yang lebih baik dibandingkan dengan *machine learning* klasik. Meski demikian, ketika dibandingkan dengan metode LSTM (*Long Short Term Memory*) sebagai model *deep learning*, hasil penelitian menunjukkan bahwa LSTM memberikan kinerja yang lebih baik daripada SVM. Hal ini menunjukkan bahwa *deep learning* memiliki keunggulan dalam memproses data kompleks seperti dalam kasus analisis sentimen, di mana LSTM dapat mengekstrak fitur yang relevan secara otomatis dari data mentah, sementara SVM membutuhkan fitur seleksi sebelumnya untuk hasil yang akurat. Maka dari itu penggunaan metode *deep learning* LSTM semakin populer dalam analisis sentimen karena kemampuannya dalam memproses data yang kompleks dan mendapatkan hasil yang lebih baik dibandingkan dengan model *machine learning* klasik [14].

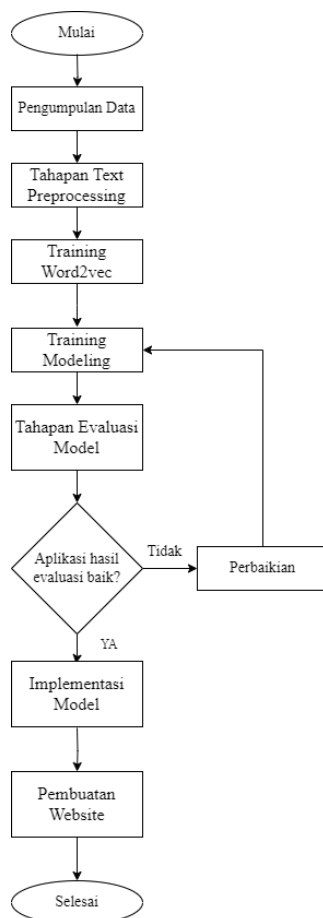
Untuk membuat analisis sentimen yang akurat, perlu dipertimbangkan metode *pretraining* yang tepat. Salah satu metode *pretraining* yang dapat digunakan adalah *Word2Vec*, karena mampu menangkap makna semantik teks dengan baik. Dalam *Word2Vec*, setiap kata yang berhubungan mempunyai vektor yang hampir mirip, sehingga dapat membantu dalam menganalisis sentimen secara lebih akurat [15][16]. Sebagai jaringan saraf yang memproses data teks, *Word2Vec* merupakan metode yang diusulkan untuk menganalisis teks. Metode ini terdiri dari dua model pembelajaran, yaitu *Continuous Bag of Words* (CBOW) dan *Skip-gram*. *Word2Vec* dapat membantu mengelompokkan kata-kata yang serupa, sehingga memudahkan dalam analisis sentimen teks [17]. Beberapa penelitian sebelumnya [18], [19] tentang klasifikasi opini publik terhadap bakal calon presiden hanya sampai pengembangan model, sehingga kebergunaan dari model yang telah dibuat tidak maksimal.

Dari permasalahan yang diuraikan, penelitian ini bertujuan untuk membuat *website* yang menampilkan *dashboard* Klasifikasi Opini Publik di Twitter Terhadap Bakal Calon Presiden Indonesia Tahun 2024 Menggunakan LSTM Secara *Realtime* Berbasis *Website*. Penelitian ini dibutuhkan untuk memudahkan bakal calon presiden Indonesia tahun 2024 dan tim kampanye dalam membaca data persebaran klasifikasi opini publik setiap provinsi secara *realtime*. Adapun pembaruan dari penelitian ini dibandingkan dengan penelitian sebelumnya adalah adanya kemampuan untuk melakukan proses klasifikasi dan visualisasi secara *realtime* pada *website* yang menampilkan *dashboard*. Manfaat dari kemampuan proses *realtime* pada

website yang menampilkan *dashboard* ini adalah memungkinkan pengguna untuk memperoleh data dan informasi yang lebih akurat dan terbaru mengenai opini publik terhadap calon presiden pada waktu yang bersamaan. Dengan demikian, penelitian ini memiliki kontribusi bagi pelaku politik (politikus pada khususnya dan masyarakat pada umumnya) dalam pengambilan keputusan berdasarkan informasi klasifikasi opini publik dari *website* yang informatif sesuai kebutuhan.

2. Metode

Penelitian ini bertujuan untuk membuat *website* yang menampilkan *dashboard* Klasifikasi Opini Publik di Twitter Terhadap Bakal Calon Presiden Indonesia Tahun 2024 Menggunakan metode LSTM Secara *Realtime* Berbasis *Website*. *Website* yang menampilkan *dashboard* ini dibutuhkan oleh bakal calon presiden Indonesia tahun 2024 dan tim kampanye untuk memudahkan membaca data persebaran klasifikasi opini publik di twitter setiap provinsi di Indonesia secara *realtime*. Metode LSTM adalah metode pembelajaran mendalam yang menawarkan akurasi lebih baik daripada metode pembelajaran mesin klasik.



Gambar 1. Prosedur Penelitian

2.1. Data Penelitian

Pada penelitian ini data yang digunakan untuk pembuatan model klasifikasi opini publik merupakan data yang bersumber dari *crawling* twitter dengan kata kunci “Ganjar Pranowo”, “Anies Baswedan”, dan “Prabowo

Subianto”. Data tersebut berupa data sekumpulan text hasil tweet dari pengguna twitter dengan rentan waktu 2020 sampai 2022.

2.2. Prosedur Penelitian

Proses penelitian melibatkan beberapa tahapan, antara lain pengumpulan data, perancangan alur program, *preprocessing* dan pre-training menggunakan metode *Word2Vec*, Tahapan *modeling*, *evaluasi model*, *implementasi model*, dan Pembuatan *Website*. Secara umum, penelitian ini mengikuti prosedur yang ditunjukkan pada Gambar 1.

2.3. Pengumpulan Data

Data yang digunakan pada penelitian ini merupakan data yang bersumber dari twitter dengan teknik pengumpulan data menggunakan *crawling* data dengan kata kunci “Ganjar Pranowo”, “Anies Baswedan”, dan “Prabowo Subianto”. Pengambilan data dilakukan dengan mengumpulkan tweet pada kurun waktu tahun 2020 sampai 2022. Data yang terkumpul selanjutnya dilabeli secara manual oleh ahli bahasa indonesia. Setelah dilakukan pelabelan data dipilih dalam jumlah yang seimbang yaitu 200 data positif, negatif, dan netral dari setiap tokoh. Sehingga menghasilkan total 1800 data tweet. Tabel 1 menampilkan hasil pelabelan data.

Tabel 1. Hasil Pelabelan Data

| Tokoh | Jumlah Data | | |
|------------------|-------------|------------|------------|
| | Positif (1) | Negatif(0) | Netral(-1) |
| Ganjar Pranowo | 200 | 200 | 200 |
| Anies Baswedan | 200 | 200 | 200 |
| Prabowo Subianto | 200 | 200 | 200 |

2.4. Text Preprocessing

Preprocessing teks bertujuan untuk mengurangi data yang tidak perlu dan mengatur data dengan cara yang membuatnya lebih mudah untuk digunakan nanti. Beberapa langkah diambil dalam fase ini, seperti menghapus data yang tidak relevan dan mengubah format data untuk memudahkan pemrosesan selanjutnya. Tahapan pada *preprocessing* antara lain:

- 1) *Case Folding*, diperlukan untuk memperoleh bentuk standar dari seluruh teks dalam suatu dokumen dengan cara mengonversinya ke dalam huruf kecil (lowercase). Tujuan dari case folding adalah membuat keseragaman dalam teks dokumen, sehingga setiap huruf dalam dokumen akan diubah menjadi huruf kecil.
- 2) *Filtering*, dalam *preprocessing* teks bertujuan untuk membersihkan data tweet dengan cara menghapus elemen-elemen yang tidak diperlukan, seperti whitespace, hastag, tanda baca, simbol, angka, html, dan mention.
- 3) Tokenisasi, adalah proses memecah teks menjadi satuan kata dengan memperhatikan karakter pemisah kata.
- 4) *Slangwords*, adalah proses untuk mengganti kata tidak baku menjadi kata baku.
- 5) *Stopword*, sebuah metode untuk mengurangi jumlah kata dengan cara menghapus kata hubung.
- 6) *Stemming*, Proses untuk menghilangkan kata imbuhan baik sufiks maupun prefiks.

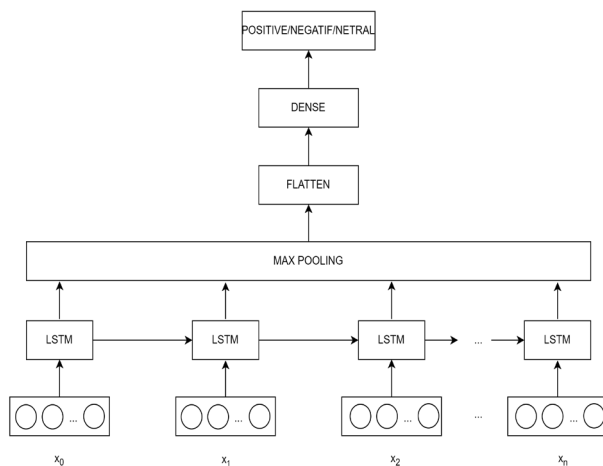
7) *Padding*, perhatikan bahwa tidak semua kalimat memiliki panjang yang sama, seperti yang dapat kita katakan secara alami beberapa kalimat panjang dan beberapa pendek. *Padding* adalah proses untuk menyamakan panjang setiap kalimat. *Padding* diperlukan karena perlu memiliki input dengan ukuran yang sama agar model dapat dibuat dan digunakan dengan baik.

2.5. Training Word2Vec

Proses pelatihan model *Word2Vec* menggunakan hasil dari penelitian[20] dengan data yang telah disesuaikan dengan data hasil *preprocessing*.

2.6. Training Modeling

Pada Penelitian ini Pembagian data dilakukan dengan membagi data menjadi 80% sebagai data latih dan 20% sebagai data uji, mengikuti *scaling law* yang ditemukan oleh Guyon pada tahun 1997[21]. *Scaling law* tersebut bertujuan untuk menentukan rasio yang optimal antara jumlah data latih dan uji dalam suatu model *machine learning*. Tujuan dari pembagian data adalah untuk memperoleh model *machine learning* yang dapat digeneralisasi dengan baik pada data baru. Rancangan struktur model terdapat pada Gambar 2.



Gambar 2. Rancangan Arsitektur LSTM dan *Word2Vec*

Pada Gambar 2 model arsitektur LSTM memproses array hasil *preprocessing* teks dengan ukuran Jumlah Kata sebagai input. Setiap kata dalam array masuk ke dalam embedding layer untuk dicarikan representasi data vektor pada array hasil *Word2Vec*. Array hasil *Word2Vec* kemudian diproses dalam LSTM masuk ke dalam pooling layer. Output dari pooling layer diolah melalui flatten layer dan dense layer dengan fungsi aktivasi softmax untuk menghasilkan output klasifikasi positif, negatif, atau netral. Tujuan dari arsitektur model LSTM ini adalah untuk menghasilkan prediksi klasifikasi pada teks yang telah diolah sebelumnya.

Sebelum masuk ke proses pelatihan dilakukan terlebih dahulu proses *Hyper Tuning Parameter* menggunakan Grid Search CV dimana parameter LSTM yang diujikan adalah *batch size*, *dropout* parameter, dan *learning rate*. Dengan

nilai parameter yang diujikan yaitu *batch size* dengan kombinasi 3, 32, 64, *dropout* dengan kombinasi 0,2; 0,5; 0,7; dan *learning rate* dengan kombinasi 0,01; 0,001; 0,0001; Tabel 2 tersaji parameter LSTM yang diuji.

Tabel 2 Parameter LSTM yang diuji

| Parameter | Nilai |
|---------------|----------------------|
| Batch Size | 3, 32, 64 |
| Dropout | 0,2; 0,5; 0,7; |
| Learning Rate | 0,01; 0,001; 0,0001; |

Proses pelatihan model LSTM bertujuan untuk menggunakan data yang telah diperoleh dan mendapatkan hasil pemodelan yang terbaik. Pada proses pelatihan, parameter-parameter yang diinisialisasi pada model LSTM seperti *batch size*, *dropout*, *learning rate*, dan parameter *Word2Vec* yang telah dilatih sebelumnya dapat mempengaruhi hasil akhir model. Dalam penelitian ini, parameter optimizer yang digunakan adalah Adam, sedangkan parameter *batch size*, *dropout*, dan *learning rate* telah diatur menggunakan *Hyper Tuning Parameter*. Tujuannya adalah untuk mendapatkan hasil pelatihan model yang optimal dan dapat menghasilkan prediksi klasifikasi yang akurat pada data uji.

2.7. Evaluasi Model

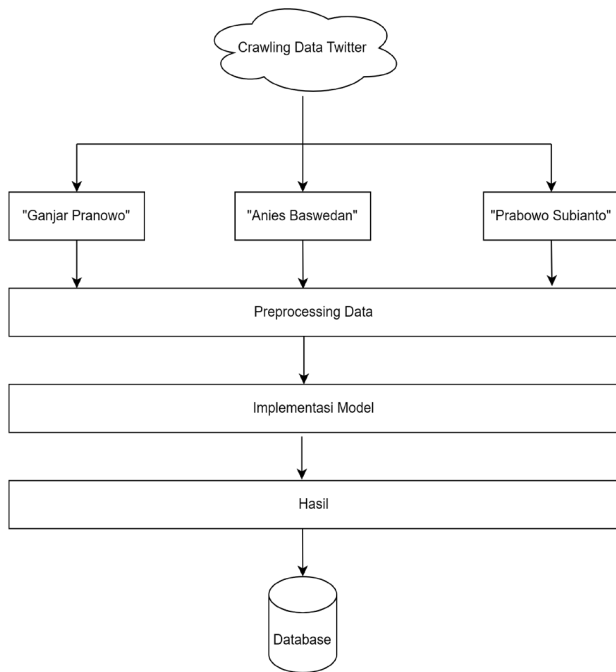
Penelitian ini mengukur kinerja model klasifikasi dengan menggunakan *classification report*. *classification report* adalah sebuah laporan yang menyediakan informasi rinci tentang kinerja model klasifikasi pada data uji. Report ini memperlihatkan beberapa nilai yang bermanfaat untuk mengevaluasi performa model klasifikasi, seperti *precision*, *recall*, *f1-score*, dan support. Tabel 3 menunjukkan contoh hasil prediksi tiga kelas yang telah dievaluasi menggunakan *classification report*.

Tabel 3 Contoh tabel hasil evaluasi

| Label | Prediksi | | |
|-----------|------------------|---------------|----------|
| | <i>precision</i> | <i>recall</i> | F1-score |
| Positif | xx | xx | xx |
| Negatives | xx | xx | xx |
| Netral | xx | xx | xx |

2.8. Implementasi Model

Implementasi model bertujuan untuk menggunakan model yang telah dibuat sebelumnya untuk data baru hasil *crawling* data twitter yang akan dilakukan klasifikasi sentimen serta memasukan data hasil klasifikasi sentimen ke dalam *database*. Gambar 3 Merupakan alur proses implementasi model.

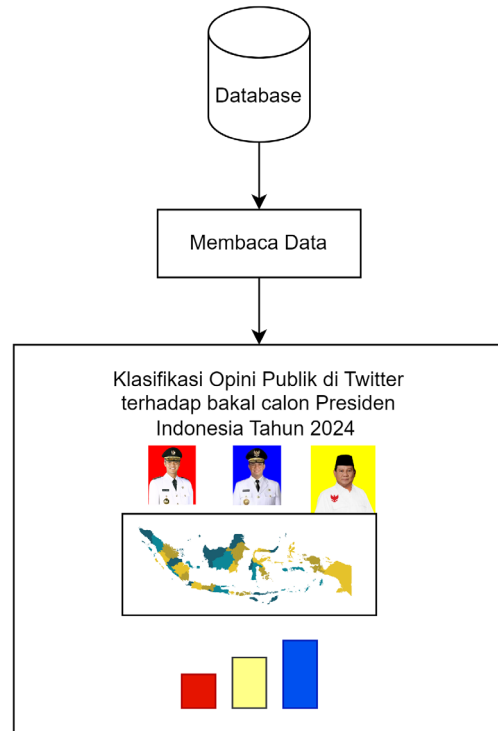


Gambar 3. Alur Implementasi Model

Proses *crawling* data twitter menggunakan library twint dengan parameter yang diisi adalah “Ganjar Pranowo”, “Anies Baswedan”, dan “Prabowo Subianto”. Parameter yang lain yaitu berupa lokasi *longitude* dan *latitude* dari setiap provinsi yang ada di Indonesia. Adapun parameter lain berupa tanggal. Setelah proses *crawling* selesai proses selanjutnya adalah *preprocessing* data hasil *crawling* dengan *preprocessing* sama seperti tahapan 2.4 *Text Preprocessing*. Setelah *preprocessing* data selesai maka dilanjutkan dengan implementasi model yang telah dibuat dengan data baru hasil *crawling* yang telah dilakukan *text preprocessing*. Proses selanjutnya adalah melakukan *store* data hasil implementasi model kedalam database. Proses pada setiap tahapannya pada penelitian ini dilakukan secara otomatis secara *realtime* setiap hari yang dijalankan menggunakan *Task Scheduler* setiap jam 00.30.

2.9. Pembuatan Website

Setelah dilakukan implementasi model tahap selanjutnya adalah pembuatan website menggunakan data hasil dari implementasi model pada database menggunakan Streamlit. Alur dalam pembuatan website diperlihatkan dalam Gambar 4.



Gambar 4. Alur Pembuatan website

3. Hasil dan Pembahasan

Penelitian ini menghasilkan suatu model LSTM dan *website* yang menampilkan *dashboard* Klasifikasi Opini Publik di twitter terhadap bakal calon presiden tahun 2024 dari hasil Implementasi model yang telah dibuat secara *Realtime* berbasis *Website*.

3.1. Hasil pengumpulan data

Pada Tabel 4 memuat contoh sampel data yang dikumpulkan dalam penelitian ini. Angka 1 merepresentasikan sentimen positif, angka 0 merepresentasikan sentimen negatif dan angka -1 merupakan sentimen netral.

Tabel 4. *Sample Dataset*

| No. | tweet | sentimen |
|-----|--|----------|
| 1. | Pak @ganjarpranowo ... Sangat memalukan sekali pendukung anda ini. Ingat pak...kualitas yang didukung tergantung pendukungnya. Pendukungnya tolong kayak gini, gimana yg didukungnya...??? | 0 |
| 2. | @aniesbaswedan Anies memang keren cadaaass 🙌🙌 https://t.co/sfCWfTU74E | 1 |
| 3. | @KalbarGerindra @Fraksi Gerindra @Gerindra @prabowo semoga situasi indonesia baik2 aja selalu. Semoga perekonomian pulih, semoga rakyat baik2 aja... | -1 |

3.2. Hasil Preprocessing

Hasil dari setiap tahap *preprocessing* dalam penelitian ini diperlihatkan dalam Tabel 5, Tabel 6, Tabel 7, Tabel 8, Tabel 9, Tabel 10, dan Tabel 11.

Tabel 5. Case *Folding*

| No. | Sebelum | Sesudah |
|-----|---|---|
| 1. | Pak @ganjarpranowo ... Sangat memalukan sekali pendukung anda ini. Ingat pak...kualitas yang didukung tergantung pendukungnya. Pendukungnya tolol kayak gini, gimana yg didukungnya...??? | pak @ganjarpranowo ... sangat memalukan sekali pendukung anda ini. ingat pak...kualitas yang didukung tergantung pendukungnya. pendukungnya tolol kayak gini, gimana yg didukungnya...??? |
| 2. | @aniesbaswedan Anies memang keren cadaaass 🙌🙌 https://t.co/sfCWfTU74E | @aniesbaswedan anies memang keren cadaaass 🙌🙌 https://t.co/sfcwf74e |
| 3. | @KalbarGerindra @Fraksi_Gerindra @Gerindra @prabowo semoga situasi indonesia baik2 aja selalu. Semoga perekonomian pulih, semoga rakyat baik2 aja... | @kalbargerindra @fraksi_gerindra @gerindra @prabowo semoga situasi indonesia baik2 aja selalu. semoga perekonomian pulih, semoga rakyat baik2 aja... |

Tabel 6. *Filtering*

| No. | Sebelum | Sesudah |
|-----|---|--|
| 1. | pak @ganjarpranowo ... sangat memalukan sekali pendukung anda ini. ingat pak...kualitas yang didukung tergantung pendukungnya. pendukungnya tolol kayak gini, gimana yg didukungnya...??? | pak ganjarpranowo sangat memalukan sekali pendukung anda ini ingat pak kualitas yang didukung tergantung pendukungnya tolol kayak gini gimana yg didukungnya |
| 2. | @aniesbaswedan anies memang keren cadaaass 🙌🙌 https://t.co/sfcwf74e | aniesbaswedan anies memang keren cadaaass 🙌🙌 |
| 3. | @kalbargerindra @fraksi_gerindra @gerindra @prabowo semoga situasi indonesia baik2 aja selalu. semoga perekonomian pulih, semoga rakyat baik2 aja... | kalbargerindra fraksi gerindra gerindra prabowo semoga situasi indonesia baik aja selalu semoga perekonomian pulih, semoga rakyat baik aja |

Tabel 7. *Tokenisasi*

| No. | Sebelum | Sesudah |
|-----|---|---|
| 1. | pak ganjarpranowo sangat memalukan sekali pendukung anda ini ingat pak kualitas yang didukung tergantung pendukungnya pendukungnya tolol kayak gini gimana yg didukungnya | ['pak', 'ganjarpranowo', 'sangat', 'memalukan', 'sekali', 'pendukung', 'anda', 'ini', 'ingat', 'pak', 'kualitas', 'yang', 'pendukung', 'pendukungnya', 'tolol', 'kayak', 'gini', 'gimana', 'yg', 'didukungnya'] |

| | | |
|----|--|---|
| | | 'didukung', 'tergantung', 'pendukungnya', 'pendukungnya', 'tolol', 'kayak', 'gini', 'gimana', 'yg', 'didukungnya'] |
| 2. | aniesbaswedan anies memang keren cadaaass 🙌🙌 | ['aniesbaswedan', 'anies', 'memang', 'keren', 'cadaaass', '🙌', '🙌'] |
| 3. | kalbargerindra fraksi gerindra gerindra prabowo semoga situasi indonesia baik aja selalu semoga perekonomian pulih, semoga rakyat baik aja | ['kalbargerindra', 'fraksi', 'gerindra', 'gerindra', 'prabowo', 'semoga', 'situasi', 'indonesia', 'baik', 'moga', 'moga', 'indonesia', 'moga', 'ekonomi', 'pulih', 'pulih', 'ekonomi', 'pulih', 'moga', 'rakyat'] |

Tabel 8. *Slangwords*

| No. | Sebelum | Sesudah |
|-----|---|--|
| 1. | ['pak', 'ganjarpranowo', 'sangat', 'memalukan', 'sekali', 'pendukung', 'anda', 'ini', 'ingat', 'pendukung', 'anda', 'ini', 'ingat', 'pak', 'kualitas', 'yang', 'didukung', 'tergantung', 'pendukungnya', 'pendukungnya', 'tolol', 'kayak', 'gini', 'gimana', 'yg', 'didukungnya'] | ['bapak', 'ganjarpranowo', 'sangat', 'memalukan', 'sekali', 'pendukung', 'anda', 'ini', 'ingat', 'bapak', 'kualitas', 'yang', 'didukung', 'tergantung', 'pendukungnya', 'pendukungnya', 'tolol', 'kayak', 'gini', 'gimana', 'yg', 'didukungnya'] |
| 2. | ['aniesbaswedan', 'anies', 'memang', 'keren', 'cadaaass', '🙌', '🙌'] | ['aniesbaswedan', 'anies', 'memang', 'keren', 'cadaaass', '🙌', '🙌', '🙌'] |
| 3. | ['kalbargerindra', 'fraksi', 'gerindra', 'gerindra', 'prabowo', 'semoga', 'situasi', 'indonesia', 'baik', 'moga', 'moga', 'indonesia', 'moga', 'ekonomi', 'pulih', 'pulih', 'moga', 'rakyat'] | ['kalbargerindra', 'fraksi', 'gerindra', 'gerindra', 'prabowo', 'semoga', 'situasi', 'indonesia', 'baik', 'saja', 'selalu', 'semoga', 'perekonomian', 'pulih', 'semoga', 'rakyat', 'baik', 'saja'], ['pertemuan', |

Tabel 9. *Stopwords*

| No. | Sebelum | Sesudah |
|-----|--|---|
| 1. | ['bapak', 'ganjarpranowo', 'sangat', 'memalukan', 'sekali', 'pendukung', 'kualitas', | ['ganjarpranowo', 'memalukan', 'pendukung', 'kualitas', |

| | | |
|----|---|---|
| | 'anda', 'ini', 'ingat', 'didukung', 'bapak', 'kualitas', 'yang', 'tergantung', 'didukung', 'tergantung', 'pendukungnya', 'pendukungnya', 'pendukungnya', 'tolol', 'pendukungnya', 'seperti', 'begini', 'didukungnya', 'bagaimana', 'yang', 'didukungnya'] | |
| 2. | ['aniesbaswedan', 'anies', 'memang', 'keren', 'cadaaass', '👍', '👍'] | ['aniesbaswedan', 'anies', 'keren', 'cadaaass', '👍', '👍'] |
| 3. | ['kalbargerindra', 'fraksi', 'gerindra', 'gerindra', 'prabowo', 'semoga', 'situasi', 'gerindra', 'indonesia', 'baik', 'saja', 'prabowo', 'selalu', 'semoga', 'semoga', 'perekonomian', 'pulihan', 'situasi', 'semoga', 'rakyat', 'baik', 'saja'], ['pertemuan', 'semoga', 'indonesia', 'perekonomian', 'pulihan', 'semoga', 'rakyat'] | |

Tabel 10. Stemming

| No. | Sebelum | Sesudah |
|-----|--|---|
| 1. | ['ganjarpranowo', 'memalukan', 'dukung', 'kualitas', 'dukung', 'gantung', 'dukung', 'dukung', 'tolol', 'dukung'] | ['ganjarpranowo', 'malu', 'dukung', 'kualitas', 'dukung', 'gantung', 'dukung', 'dukung', 'tolol', 'dukung'] |
| 2. | ['aniesbaswedan', 'anies', 'keren', 'cadaaass', '👍', '👍'] | ['aniesbaswedan', 'anies', 'keren', 'cadaaass'] |
| 3. | ['kalbargerindra', 'fraksi', 'gerindra', 'gerindra', 'prabowo', 'semoga', 'situasi', 'gerindra', 'indonesia', 'baik', 'saja', 'prabowo', 'selalu', 'semoga', 'semoga', 'perekonomian', 'pulihan', 'situasi', 'semoga', 'rakyat'] | ['kalbargerindra', 'fraksi', 'gerindra', 'gerindra', 'prabowo', 'moga', 'situasi', 'indonesia', 'moga', 'ekonomi', 'pulihan', 'moga', 'rakyat'] |

Tabel 11. Padding

| No. | Sebelum | Sesudah |
|-----|---|---|
| 1. | ['ganjarpranowo', 'malu', 'dukung', 'kualitas', 'dukung', 'gantung', 'dukung', 'dukung', 'tolol', 'dukung'] | [0, 0, 0, ..., 1045, 553, 790] |
| 2. | ['aniesbaswedan', 'anies', 'keren', 'cadaaass'] | [0, 0, 0, ..., 3, 101, 35, 40, 2, 6, 16, 12, |

| | |
|----|---|
| | 13, 9, 15, 14, 17, 19, 21] |
| 3. | ['kalbargerindra', 'fraksi', 'gerindra', 'gerindra', 'prabowo', 'semoga', 'situasi', 'indonesia', 'moga', 'ekonomi', 'pulihan', 'moga', 'rakyat'] |

3.3. Hasil Pretraining Menggunakan Word2Vec

Pretraining dengan menggunakan metode Word2Vec menghasilkan vektor untuk setiap kata yang terdapat pada data. Sebagai contoh, pada Gambar 5, terlihat hasil vektor yang merepresentasikan kata "keren" dengan dimensi vektor sebesar 200.

```
[ ] word_vectors.wv['keren']
array([-3.83591533e-01, 3.29930365e-01, 9.21615124e-01, 8.47023726e-03,
1.05347085e+00, 5.07972121e-01, 8.6037257e-01, 2.87288964e-01,
-1.59404725e-02, 5.72606027e-01, -5.19606650e-01, 7.39738703e-01,
-9.49131399e-02, 9.03081417e-01, -6.84528947e-01, 6.21681631e-01,
-1.01786625e+00, -7.97688127e-01, -1.84705436e-01, -2.9252240e-01,
1.43788052e+00, -4.91840988e-01, 4.55642372e-01, -1.19070402e+00,
-2.03535699e-01, 1.32962208e+00, 0.32413936e-01, -2.67046094e-01,
-1.02907785e+00, -6.38076246e-01, -5.09872198e-01, 1.08244562e+00,
2.26450756e-01, 9.75139439e-01, -1.11884785e+00, -4.83161420e-01,
-8.04402769e-01, 7.59765729e-02, -2.56538218e+00, -5.18565893e-01,
-4.33504909e-01, 2.87172824e-01, -5.41513324e-01, 1.33961308e+00,
-4.22258914e-01, 2.50645304e+00, 4.81203794e-01, -8.16274345e-01,
-6.51880383e-01, 3.02422434e-01, 8.68690386e-02, -2.79279828e-01,
-1.10343434e-01, 2.13069391e+00, 8.05422604e-01, 4.43106890e-01,
2.79550970e-01, 1.61275099e+00, -1.02223074e+00, -1.30959098e+00,
-1.26839852e+00, -4.90892128e-01, 3.88047409e-01, 3.21273633e-01,
4.26779777e-01, 8.23859930e-01, 4.97317314e-01, 8.53712857e-01,
5.78113794e-01, -1.08932543e+00, 8.84309888e-01, -1.32541156e+00,
-8.31761301e-01, 3.90070975e-02, -8.22660565e-01, -1.03608513e+00,
-7.50383019e-01, 5.35237901e-02, 1.02265990e+00, -8.99191737e-01,
4.88596946e-01, 4.62985437e-01, 9.32987481e-02, 3.84227298e-02,
-2.32710525e-01, 5.69611005e-02, -1.23269260e+00, -6.8235523e-01,
4.16893274e-01, -4.66899574e-01, 1.64221156e+00, 6.52420938e-01,
-1.54665890e+00, 1.45103551e+00, -1.43773210e+00, 2.57272720e-01,
1.31501400e+00, -1.92323327e+00, 1.25115919e+00, -1.10866323e+00,
-5.17130136e-01, 9.53280509e-01, -1.04986882e+00, 1.28013515e+00,
1.81446481e+00, 5.02952486e-01, -1.59687459e+00, 5.41543774e-02,
-2.89103240e-01, 8.06112766e-01, 1.14268327e+00, -2.50769138e-01,
4.84969079e-01, 1.82883096e+00, -2.32479826e-01, -3.43287945e-01,
-2.48253584e-01, -8.68914723e-01, -6.01742804e-01, -3.28824192e-01,
-3.19902301e-02, 1.40008008e+00, 1.13750421e-01, 7.50149548e-01,
-1.51608093e+00, -5.58978379e-01, 2.61149573e+00, 2.20504100e-01,
-3.66436541e-01, 1.06280174e+00, 4.53997012e-01, -4.28270425e-01,
-4.79455766e-01, 5.17654777e-01, -1.46101162e-01, -5.30696097e-01,
2.42267892e-01, 3.01667008e-01, 7.81534553e-01, -3.86865944e-01,
4.82040882e-01, 1.04878557e+00, -1.50074720e+00, -1.28171011e-03,
1.44753015e+00, -5.21217659e-02, 7.86526978e-01, -5.79817414e-01,
-2.33658999e-01, -3.86820585e-01, 3.49361479e-01, -1.34947538e+00,
-8.05598736e-01, 2.12883449e+00, -1.07504524e-01, 1.77587581e+00,
-1.21560764e+00, -5.0531337e-01, -2.08049250e+00, 4.34565365e-01,
-1.74039389e+00, -7.16435909e-01, 7.75245130e-01, -2.70736651e-01,
6.71027950e-01, -6.36152656e-01, 6.78133932e-01, 2.03993474e-01,
3.56262207e-01, -1.37136090e+00, 5.22586823e-01, -1.31162691e+00,
5.34358799e-01, -3.34686100e-01, 3.06720757e+00, 4.94948149e-01,
7.00747490e-01, 7.80662000e-01, -1.27233696e+00, 2.74892002e-01,
4.05646920e-01, -1.55539364e-01, -1.27046072e+00, 6.3152192e-01,
-1.3396137e+00, 1.94791079e-01, -4.24514443e-01, -2.18456697e+00,
3.18040699e-01, -1.94690764e+00, -1.69659984e+00, -1.39561655e+00,
4.93700266e-01, -1.17115533e+00, 4.71981019e-02, 9.85419452e-02,
3.86376917e-01, 9.02682900e-01, -1.96270621e+00, -3.90060494e-01],
dtype='float32')
```

Gambar 5. Word2vec

Hypertunning Parameter menggunakan GridsearchCV menghasilkan nilai parameter terbaik adalah batch size 32, dropout 0.5, dan learning rate 0.01.

3.4. Evaluasi Matrik

Hasil dari evaluasi matrik menggunakan classification report menunjukkan hasil akurasi sebesar 76%, serta matriks precision, recall dan f1-score yang dijelaskan pada Tabel 12.

Tabel 12 Hasil Evaluasi Metrik

| Label | Prediksi | | |
|-----------|-----------|--------|----------|
| | precision | recall | f1-score |
| Positif | 0,83 | 0,76 | 0,79 |
| Negatives | 0,75 | 0,83 | 0,79 |
| Netral | 0,67 | 0,67 | 0,67 |

3.5. Hasil Implementasi Model

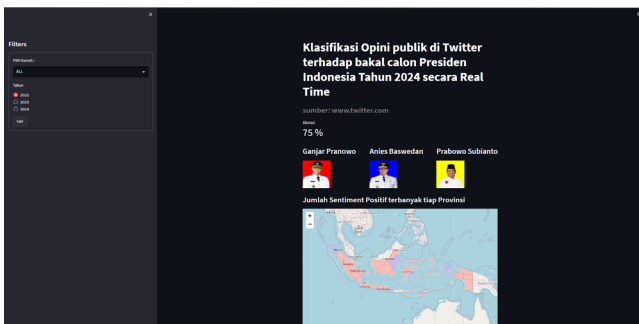
Langkah selanjutnya yaitu mengimplementasikan model yang telah dibuat dengan data *realtime crawling* dari twitter kemudian data tersebut dilakukan *preprocessing* lalu di *testing* menggunakan model yang telah dibuat setelah itu hasilnya dimasukkan kedalam database. Gambar 6 merupakan tampilan data Hasil yang telah dimasukkan kedalam database

| date | username | tweet | location | Tokoh | Sentiment |
|------------|-----------------|--|----------|----------------|-----------|
| 2023-01-16 | Hendrik66752917 | @ganjarpranowo Masih jaman PENGICTRAAN bung. Keliat... | ACEH | Ganjar Pranowo | 0 |
| 2023-01-16 | irvan_yonas | @RamiRizal Setidaknya apa yang disampaikan oleh P... | ACEH | Ganjar Pranowo | -1 |
| 2023-01-16 | Aneff92839705 | @0512Vip @AndiSinulingga @ganjarpranowo Kamu iri y... | ACEH | Ganjar Pranowo | -1 |
| 2023-01-16 | Rey_Annd | @jakapradipta @Inwan2yah1 @DivHumas_Polri @ganjarp... | ACEH | Ganjar Pranowo | 0 |
| 2023-01-16 | makganjarsumut | Mak Ganjar Sumatera Utara berbagi kebaikan. #m... | ACEH | Ganjar Pranowo | -1 |
| 2023-01-16 | Block13_Madura | @AlimDatangLagi @ganjarpranowo. Jangan gotong royo... | ACEH | Ganjar Pranowo | -1 |
| 2023-01-15 | ND45MUpetal | @ganjarpranowo Gap WADAS piye pak? Namung tanggile... | ACEH | Ganjar Pranowo | -1 |
| 2023-01-15 | SHBcrypto | @inwan2yah1 @DivHumas_Polri @ganjarpranowo Gass @D... | ACEH | Ganjar Pranowo | -1 |
| 2023-01-15 | Daniel07568575 | Po'tren'si Kemajuan Pak Ganjar Sangat Besar #ganja... | ACEH | Ganjar Pranowo | 1 |
| 2023-01-15 | Daniel07568575 | Gue sih setuju banget kalo pak ganjar cocok jadi L... | ACEH | Ganjar Pranowo | 1 |
| 2023-01-15 | Daniel07568575 | Pak ganjar bisa kerja untuk Indonesia di bandingka... | ACEH | Ganjar Pranowo | 1 |
| 2023-01-15 | Daniel07568575 | Pak Ganjar bakal jadi Presiden Indonesia 2024 #gan... | ACEH | Ganjar Pranowo | 1 |
| 2023-01-15 | tribunmedan | Guntur Soekarnoputra Blakblakan Pih Ganjar Prano... | ACEH | Ganjar Pranowo | -1 |
| 2023-01-15 | FandikaDimas | kinerja Pak Ganjar sangat hebat #ganjarpranowo #g... | ACEH | Ganjar Pranowo | 1 |

Gambar 6. Database

3.6. Hasil Pembuatan Website

Langkah selanjutnya adalah membuat *website* yang menampilkan *dashboard* Visualisasi Klasifikasi Opini Publik dengan menggunakan data dari Database. *website* yang menampilkan *dashboard* ini memiliki beberapa fitur, diantaranya Gambar 7 yang memungkinkan menampilkan peta warna berdasarkan tingkat sentimen positif yang paling banyak ditemukan pada bakal calon presiden dari setiap provinsi, serta Gambar 8 yang menampilkan jumlah keseluruhan klasifikasi sentimen setiap bakal calon presiden secara *realtime*. Selain itu, *website* yang menampilkan *dashboard* ini juga dilengkapi dengan fitur untuk melakukan filter terhadap hasil klasifikasi sentimen dari setiap provinsi dan tahun tertentu.



Gambar 7. Dashboard 1



Gambar 8. Dashboard 2

Proses *hypertunning* menggunakan *gridsearchCV* dengan parameter pada model LSTM dengan menggunakan nilai parameter *batch size*, *dropout*, dan *learning rate*. Nilai-nilai parameter yang diujikan pada proses *hypertunning* sebagai berikut :

- 1) *batch size* dengan kombinasi 3, 32, 64
- 2) *dropout* dengan kombinasi 0.2, 0.5, 0.7
- 3) *learning rate* dengan kombinasi 0.01, 0.001.

Proses *hypertunning* dilakukan dengan tujuan untuk mencari kombinasi nilai parameter terbaik yang dapat menghasilkan akurasi prediksi yang tinggi. Setelah dilakukan proses *hypertunning*, ditemukan bahwa nilai parameter terbaik adalah *batch size* 32, *dropout* 0.5, dan *learning rate* 0.01. Tabel 13 merupakan detail Hasil *hypertunning* parameter dengan *GridsearchCV*

Tabel 13. Hasil Hypertunning Parameter

| No. | Batch Size | Dropout Rate | Learning Rate | Akurasi |
|-----|------------|--------------|---------------|---------|
| 1 | 3 | 0.2 | 0.01 | 0.17624 |
| 2 | 3 | 0.2 | 0.001 | 0.10444 |
| 3 | 3 | 0.2 | 0.0001 | 0.26762 |
| 4 | 3 | 0.5 | 0.01 | 0.57441 |
| 5 | 3 | 0.5 | 0.001 | 0.15666 |
| 6 | 3 | 0.5 | 0.0001 | 0.39164 |
| 7 | 3 | 0.7 | 0.01 | 0.26110 |
| 8 | 3 | 0.7 | 0.001 | 0.05875 |
| 9 | 3 | 0.7 | 0.0001 | 0.18930 |
| 10 | 32 | 0.2 | 0.01 | 0.33290 |
| 11 | 32 | 0.2 | 0.001 | 0.03264 |
| 12 | 32 | 0.2 | 0.0001 | 0.56136 |
| 13 | 32 | 0.5 | 0.01 | 0.76054 |
| 14 | 32 | 0.5 | 0.001 | 0.63107 |
| 15 | 32 | 0.5 | 0.0001 | 0.13055 |
| 16 | 32 | 0.7 | 0.01 | 0.14360 |
| 17 | 32 | 0.7 | 0.001 | 0.45692 |
| 18 | 32 | 0.7 | 0.0001 | 0.05875 |
| 19 | 64 | 0.2 | 0.01 | 0.27415 |
| 20 | 64 | 0.2 | 0.001 | 0.70496 |
| 21 | 64 | 0.2 | 0.0001 | 0.07833 |
| 22 | 64 | 0.5 | 0.01 | 0.49608 |
| 23 | 64 | 0.5 | 0.001 | 0.63995 |
| 24 | 64 | 0.5 | 0.0001 | 0.01305 |
| 25 | 64 | 0.7 | 0.01 | 0.50261 |
| 26 | 64 | 0.7 | 0.001 | 0.00522 |
| 27 | 64 | 0.7 | 0.0001 | 0.67001 |

Dengan menggunakan kombinasi nilai parameter terbaik, model LSTM mampu menghasilkan akurasi prediksi sebesar 76%. Selanjutnya, model LSTM digunakan untuk membuat sebuah *website* yang menampilkan *dashboard* yang dapat melakukan klasifikasi sentimen terhadap data tweet yang berisi komentar terkait bakal calon presiden. *Dashboard* ini memiliki beberapa fitur, di antaranya adalah menampilkan peta warna berdasarkan tingkat sentimen positif yang paling banyak ditemukan pada bakal calon presiden dari setiap provinsi, menampilkan jumlah keseluruhan klasifikasi sentimen setiap bakal calon presiden, dan melakukan filter terhadap hasil klasifikasi sentimen dari setiap provinsi dan tahun tertentu.

Fitur peta warna sentimen berdasarkan provinsi dapat digunakan untuk melihat sebaran tingkat sentimen positif terhadap setiap bakal calon presiden dari setiap provinsi. Peta warna ini dibuat dengan memanfaatkan data hasil klasifikasi sentimen yang dilakukan oleh model LSTM. Selain itu, *website* yang menampilkan *dashboard* ini juga dapat menampilkan jumlah keseluruhan klasifikasi sentimen setiap bakal calon presiden, sehingga pengguna dapat melihat popularitas setiap bakal calon presiden berdasarkan jumlah klasifikasi sentimen.

Selain itu, *website* yang menampilkan *dashboard* ini juga dilengkapi dengan fitur filter hasil klasifikasi sentimen dari setiap provinsi dan tahun tertentu. Fitur ini memudahkan pengguna untuk melakukan analisis sentimen pada data tweet secara lebih terperinci. Pengguna dapat melakukan filter berdasarkan provinsi dan tahun tertentu untuk melihat hasil klasifikasi sentimen yang spesifik.

4. Kesimpulan

Hasil penelitian mengenai klasifikasi opini publik terhadap bakal calon presiden Indonesia tahun 2024 menggunakan model LSTM menunjukkan bahwa model tersebut memiliki akurasi sebesar 76%, serta nilai *precision*, *recall*, dan F1-score masing-masing sebesar 75%. Selain itu, hasil *hypertuning* parameter menggunakan GridSearchCV menunjukkan bahwa *batch size* yang optimal adalah 32, *learning rate* yang optimal adalah 0.01, dan *dropout* yang optimal adalah 0.5. Implementasi model klasifikasi opini publik terhadap bakal calon presiden Indonesia tahun 2024 telah dilakukan dengan menyimpan hasil implementasi model ke dalam database. Selain itu, proses *crawling* data dan storing data hasil implementasi model dilakukan secara *realtime* menggunakan Task Scheduler. Proses tersebut dijadwalkan untuk berjalan setiap hari pada pukul 00.30 dini hari.

Ide untuk penelitian selanjutnya adalah melakukan perbandingan dengan penelitian ini menggunakan teknik *crawling* yang berbeda seperti Selenium, dan juga menggunakan metode *deep learning* yang berbeda, teknik *realtime* yang berbeda, dan cara pembuatan *website* yang menampilkan *dashboard* yang berbeda. Tujuannya adalah untuk memperluas wawasan dan memperoleh hasil yang lebih komprehensif.

Daftar Pustaka

- [1] Sekretariat Negara, *Undang - Undang Republik Indonesia Nomor 42 Tahun 2008 Tentang Pemilihan Umum Presiden dan Wakil Presiden*. 2008.
- [2] Y. Medistiara, "Survei PWS: Elektabilitas Prabowo 30,8%, Ganjar 18,8%, Anies 17,5%," Jakarta, 2022.
- [3] N. Utami, "Survei Capres IPO: Prabowo 24,8%, Anies 22,5%, dan Ganjar 19,3%," Jakarta, 2022.
- [4] M. O. Erwanti, "Survei Capres SMRC: Ganjar 32,1%, Prabowo 27,5%, Anies 26%," Jakarta, 2022.
- [5] A. B. Ramadhan, "Survei Charta Politika: Prabowo Unggul di Jabar, Ganjar di Sumut dan Kaltim," Jakarta, 2022.
- [6] D. Reportal, "Social media statistics for Indonesia in 2022." 2022. <https://datareportal.com/reports/digital-2022-indonesia> (accessed Nov. 03, 2022).
- [7] N. I. Widiastuti, E. Rainarli, and K. E. Dewi, "Peringkasan dan Support Vector Machine pada Klasifikasi Dokumen," *J. Infotel*, vol. 9, no. 4, p. 416, 2017, doi: 10.20895/infotel.v9i4.312.
- [8] D. C. Mutz, S. American, P. Science, and N. Oct, "The Consequences of Cross-Cutting Networks for Political Participation Author (s): Diana C . Mutz Reviewed work (s): Source : American Journal of Political Science , Vol . 46 , No . 4 (Oct . , 2002) , pp . 838-855 Published by : Midwest Political Scie.," vol. 46, no. 4, pp. 838–855, 2013.
- [9] B. Liu, *Sentiment Analysis: A Fascinating Problem*. 2012.
- [10] T. Iqbal and S. Qureshi, "The survey: Text generation models in *deep learning*," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 6, pp. 2515–2528, 2022, doi: 10.1016/j.jksuci.2020.04.001.
- [11] A. Yenter and A. Verma, "Deep CNN-LSTM with combined kernels from multiple branches for IMDB review sentiment analysis," *2017 IEEE 8th Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf. UEMCON 2017*, vol. 2018-Janua, pp. 540–546, 2017, doi: 10.1109/UEMCON.2017.8249013.
- [12] A. P. Nardilasari, A. L. Hananto, S. S. Hilabi, and B. Priyatna, "Analisis Sentimen Calon Presiden 2024 Menggunakan Algoritma SVM," vol. 7, no. 1, pp. 11–18, 2024.
- [13] S. Wiyono, D. S. Wibowo, M. F. Hidayatullah, and D. Dairoh, "Comparative Study of KNN, SVM and Decision Tree Algorithm for Student's Performance Prediction," *Int. J. Comput. Sci. Appl. Math.*, vol. 6, no. 2, p. 50, 2020, doi: 10.12962/j24775401.v6i2.4360.
- [14] H. Ghulam, F. Zeng, W. Li, and Y. Xiao, "Deep Learning-Based Sentiment Analysis for Roman Urdu Text," *Procedia Comput. Sci.*, vol. 147, pp. 131–135, 2019, doi: 10.1016/j.procs.2019.01.202.
- [15] R. P. Nawangsari, R. Kusumaningrum, and A. Wibowo, "Word2vec for Indonesian sentiment analysis towards hotel reviews: An evaluation study," *Procedia Comput. Sci.*, vol. 157, pp. 360–366, 2019, doi: 10.1016/j.procs.2019.08.178.
- [16] D. Jatnika, M. A. Bijaksana, and A. A. Suryani, "Word2vec model analysis for semantic similarities in English words," *Procedia Comput. Sci.*, vol. 157, pp. 160–167, 2019, doi: 10.1016/j.procs.2019.08.153.
- [17] C. Zhang, X. Wang, S. Yu, and Y. Wang, "Research on Keyword Extraction of Word2vec Model in Chinese Corpus," *Proc. - 17th IEEE/ACIS Int. Conf. Comput. Inf. Sci. ICIS 2018*, pp. 339–343, 2018, doi: 10.1109/ICIS.2018.8466534.
- [18] E. Miranda, "Sentiment Analysis using Sentiwordnet and Machine Learning Approach (Indonesia general election opinion from the twitter content)," *2019 Int. Conf. Inf. Manag. Technol.*, vol. 1, no. August, pp. 62–67, 2019.
- [19] P. Singh, R. S. Sawhney, and K. S. Kahlon, "Forecasting the 2016 US Presidential Elections Using Sentiment Analysis," vol. 1, pp. 276–288, 2017, doi: 10.1007/978-3-319-68557-1_36.
- [20] D. I. Af'idah, D. Dairoh, S. F. Handayani, R. W. Pratiwi, and S. I. Sari, "Sentimen Ulasan Destinasi Wisata Pulau Bali Menggunakan Bidirectional Long Short Term Memory," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 3, pp. 607–618, 2022, doi: 10.30812/matrik.v21i3.1402.
- [21] I. M. Guyon, "A Scaling Law for the Validation-Set Training-Set Size Ratio," 1997.