

Analisis Sentimen Program Sekolah Rakyat Menggunakan Metode Bidirectional Long Short Term Memory

Sentiment Analysis of Sekolah Rakyat Program Using Bidirectional LSTM

Fina Qurrotuna^{*1}, Karina Auliasari², Yosep Agus Pranoto³

^{1,2,3}Institut Teknologi Nasional Malang; Jln. Raya Karanglo Kampus.2 Malang, Jawa Timur,
Indonesia, +62-341-417636

^{1,2,3}Jurusan Teknik Informatika, FTI ITN, Malang

e-mail: ¹qurrotunafina1@gmail.com, ²karina.auliasari@lecturer.itn.ac.id,
³yoa@lecturer.itn.ac.id

Abstrak

Sekolah rakyat yang berada di bawah kementerian sosial merupakan sekolah berbasis asrama yang ditujukan bagi anak sekolah yang berasal dari keluarga miskin dan miskin ekstrem. Sekolah ini bertujuan memberikan pendidikan formal dan non formal, menjadikan peserta didik menjadi lulusan yang cerdas dan berketerampilan, dan harapannya dapat memperbaiki kehidupan dirinya dan keluarganya. Penelitian ini ditujukan untuk melakukan analisis sentimen terhadap opini masyarakat seputar program pemerintah yang ada pada bidang pendidikan yaitu sekolah rakyat dengan menggunakan metode deep learning yaitu metode Bidirectional Long Short Term Memory (BiLSTM). Data yang digunakan dalam penelitian ini didapat dari platform X (x.com) dengan jumlah 4714. Dari hasil penelitian yang telah dilakukan inilah kemudian didapatkan informasi bahwa mayoritas sentimen masyarakat mengenai program sekolah rakyat adalah sentimen positif dengan total 3683 atau sebanyak 78.1%. Model BiLSTM berhasil melakukan proses klasifikasi dengan baik dibuktikan dengan nilai akurasi, presisi dan recall yang mencapai 89%. Temuan ini dapat memberikan pengetahuan tentang bagaimana opini masyarakat terkait program sekolah rakyat dan dapat membantu dalam evaluasi kebijakan program

Kata kunci—3-5 analisis sentimen, sekolah rakyat, bidirectional long-short term memory

Abstract

The Sekolah Rakyat program under the Ministry of Social Affairs are boarding schools for children from poor and extremely poor families. These program intend to provide formal and informal education, so that transforming students into intelligent and skilled graduates, and hopefully improving the lives of themselves and their families and not continuously live in a cycle of poverty. This study intend to analyze public opinion on the sekolah rakyat program using the Bidirectional Long Short Term Memory (BiLSTM) method. The data used in this study were obtained from the X platform (x.com) with a total of 4714. From this study, information was obtained that the majority of sentiment was positive with a total of 3683 or 78.1%. The BiLSTM model developed and used in this study successfully recognized all classes, both positive and negative, and carried out the classification process well. This was proven by the performance matrix values obtained, namely accuracy, precision, and recall values reaching 89%. These

findings can provide insight into how public opinion is related to the sekolah rakyat program and can assist in evaluating program policies.

Keywords—3-5 sentiment analysis, sekolah rakyat, bidirectional long-short term memory

1. PENDAHULUAN

Pemerintahan Prabowo Gibran genap satu tahun pada tanggal 24 Oktober 2025 lalu (Setpres, 2024). Setahun pemerintahan ini banyak melahirkan program baru salah satunya program sekolah rakyat. Sekolah Rakyat merupakan sekolah berbasis asrama yang bertujuan memberikan pendidikan formal dan non formal, menjadikan peserta didik menjadi lulusan yang cerdas dan berketerampilan, dan harapannya dapat memperbaiki kehidupan dirinya dan keluarganya (Noerdin, 2025). Sekolah ini diperuntukkan secara khusus bagi anak-anak usis sekolah yang berasal dari keluarga miskin dan miskin ekstrem yang selama ini kesulitan terhadap akses pendidikan berkualitas sehingga menyebabkan mereka putus sekolah bahkan belum pernah sama sekali mengenyam pendidikan.

Sejak diresmikan langsung oleh bapak Presiden Prabowo Subianto pada 14 Juli 2025 lalu, program sekolah rakyat ini banyak sekali memunculkan pro dan kontra dikalangan masyarakat. Perbedaan pendapat ini kemudian disampaikan mereka salah satunya melalui platform media sosial X. Salah satu cara efektif untuk memahami opini masyarakat adalah dengan melakukan analisis sentimen. Dalam penelitian yang akan dilakukan nantinya proses analisis sentimen akan dikembangkan dengan menggunakan metode BiLSTM. Dimana data sentimen masyarakat mengenai program sekolah rakyat yang diambil dari platform X ini nantinya akan diklasifikasikan oleh model menjadi 2 kelas yaitu kelas positif dan kelas negatif.

Penelitian yang dilakukan oleh Dimas Thaqif Attaulah dan Dewi Soyusiawaty terkait analisis sentimen menggunakan metode Bi LSTM menunjukkan bahwa metode ini mampu mencapai akurasi sebesar 88,75%, dengan presisi 88,9%, *recall* 88,8%, dan *f1-score* 88,8%. Namun topik pada penelitian ini hanya terbatas pada program amakan bergizi gratis saja (Attaulah & Soyusiawaty, 2025).

Belum banyak penelitian yang membahas analisis sentimen terkait program sekolah rakyat ini, dikarenakan program ini merupakan program yang tergolong masih baru. Maka dari itu saya merasa sangat perlu untuk melakukan proses analisis guna melihat sejauh mana pandangan masyarakat mengenai program sekolah rakyat yang sudah berjalan ini dan bagaimana tanggapan masyarakat mengenai program sekolah rakyat ini. Dimana hasil dari analisis ini nantinya dapat dimanfaatkan sebagai bahan evaluasi program maupun pembuatan kebijakan baru terkait program sekolah rakyat.dimasukkan dalam bagian ini.

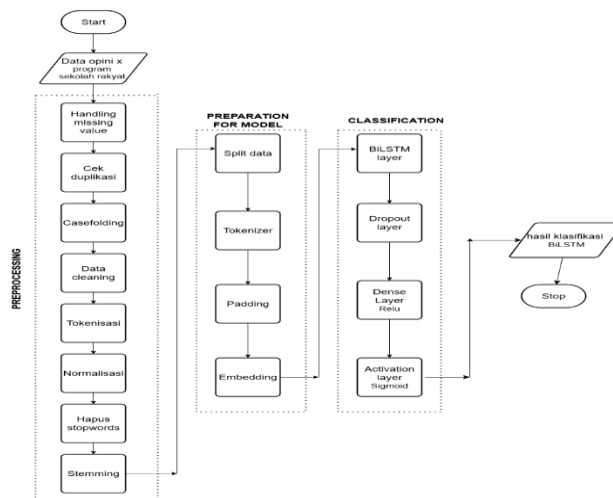
2. METODE PENELITIAN

2.1 Analisis Sistem

Sistem yang dikembangkan merupakan sistem analisis sentimen berbasis website yang digunakan untuk mengklasifikasikan sentimen masyarakat mengenai program sekolah rakyat menjadi sentimen negatif atau positif. Sistem website analisis sentimen ini akan dirancang dan dikembangkan dengan menggunakan bahasa pemrograman Python versi 3.11.9 dan kerangka kerja Streamlit versi 1.50 yang akan diintegrasikan dengan menggunakan database mongoDB dengan menggunakan modul pymongo.

Analisis Sentimen Program Sekolah Rakyat Menggunakan Metode Bidirectional Long Short Term Memory

2.2 Flowchart Sistem



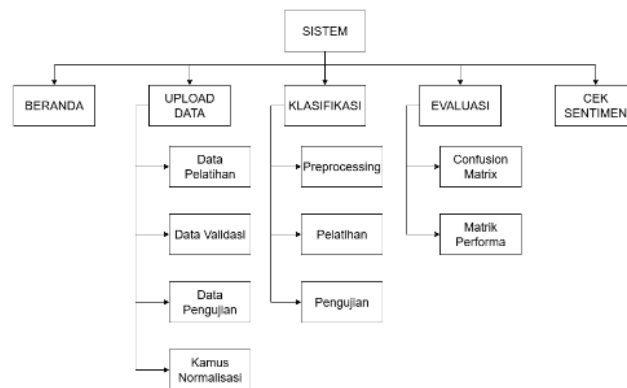
Gambar 1. Flowchart sistem

Gambar 1. merupakan alur sistem analisis sentiment pada tahap preprocessing meliputi meliputi *handling missing value*, cek duplikasi data, data *cleaning*, normalisasi, tokenisasi, hapus *stopwords*, dan *stemming*. Setelah semua tahapan yang ada pada proses *preprocessing* atau prapemrosesan selesai dilakukan, selanjutnya data akan masuk ke proses pembagian dataset. Dataset yang akan digunakan dalam sistem website analisis sentimen yang akan dikembangkan nantinya akan dibagi menjadi tiga buah file data, yaitu data pelatihan (*data training*), data validasi dan data pengujian (*data testing*).

Data pelatihan (*data training*) dan data validasi inilah yang nantinya akan digunakan dalam sistem analisis sentimen saat proses pelatihan model atau proses *training* sedang berlangsung. Sedangkan data pengujian akan digunakan saat proses pengujian model atau proses *testing*. Setelah tahapan proses pembagian dataset selesai dilakukan kemudian dilanjutkan dengan tahapan proses tokenisasi, dimana pada tahapan tokenisasi inilah yang akan menjadikan teks menjadi representasi numerik atau angka. Jadi data teks yang awalnya berupa huruf akan dikodekan atau diubah menjadi angka dalam proses tokenisasi ini, proses ini dilakukan dengan menggunakan tokenizer. Setelah tokenisasi selesai maka data akan masuk pada tahap *padding*. Proses *padding* ditujukan untuk menyamakan panjang dimensi data, jadi data yang awalnya berbeda panjang akan disamaratakan agar dapat diproses oleh model nantinya. Setelah proses *padding* selesai maka hasilnya seluruh panjang data akan sama dan kemudian barulah masuk pada proses *embedding*. Proses *embedding* ini akan memetakan data numerik tadi kedalam representasi vektor multidimensi. Setelah proses *padding* selesai selanjutnya data masuk ke tahapan klasifikasi.

Pada proses klasifikasi ini data akan masuk proses dimana pertama akan masuk pada layer Bi LSTM, selanjutnya masuk ke *layer Dropout*, *layer* ini berfungsi untuk mencegah overfitting dengan cara mematikan beberapa *neuron* secara acak saat proses *training* berlangsung. Selanjutnya masuk pada *Dense layer* atau *Hidden layer*, biasanya pada *layer* ini digunakan fungsi aktivasi seperti *ReLU* yang berfungsi untuk mengubah representasi dari LSTM menjadi fitur yang lebih padat sebelum masuk ke *output*. *Layer* terakhir yaitu *layer output*, pada *layer* ini fungsi aktivasi yang digunakan sesuai dengan jumlah kelas yang dikehendaki saat klasifikasi, jika hanya 2 kelas bisa menggunakan fungsi aktivasi *sigmoid*, sedangkan untuk 3 kelas bisa menggunakan fungsi aktivasi *softmax*.

2.3 Struktur Menu



Gambar 2. Struktur menu

Dapat dilihat pada gambar 2 dalam struktur menu website analisis sentimen terdapat beberapa menu yaitu ada Beranda, dimana di halaman ini akan menampilkan jumlah sentimen positif dan juga negatif, visualisasi sentimen positif dan negatif menggunakan *pie chart*, dan akurasi. Selanjutnya ada menu Upload Data, pada menu ini pengguna diharuskan mengupload 4 buah file data berformat .csv yaitu data pealtihan, data validasi, data pengujian dan kamus normalisasi. Selanjutnya ada menu Klasifikasi, pada menu dilakukan proses *preprocessing* data sampai dengan proses klasifikasi. Selanjutnya terdapat menuEvaluasi data yang menampilkan hasil pengujian pada data pengujian, *confusion matrix*, dan matrik performa model. Terakhir adalah menu Cek Sentimen, pada halaman ini pengguna dapat memasukkan sentimen kemudian mengklasifikasikannya secara langsung dan hasilnya akan ditampilkan beserta tingkat akurasi

2.4 Bidirectional Long Short Term Memory (BiLSTM)

Metode BiLSTM merupakan hasil pengembangan dari metode LSTM dimana metode BiLSTM ini memiliki dua lapis LSTM yang bekerja dari dua arah yang berbeda, yaitu *forward* dan *backward*. *Layer forward* adalah bagian dari BiLSTM dimana layer ini yang akan bekerja dengan cara memproses kalimat secara maju dari arah kiri ke kanan, sedangkan *layer bakcward* adalah bagian dari BiLSTM dinamakan layer inilah yang akan bekerja dengan cara memproses kalimat secara maju dari arah kanan ke kiri. Hal ini memungkinkan BiLSTM menangkap konteks kalimat secara keseluruhan dari masa lalu (*forward*) dan masa depan (*backward*) (Bouchra, Suarjaya, & Rusjyanthi, 2024).

Persamaan LSTM per arah

1. Forget Gate

Berfungsi untuk mengontrol informasi dari proses sebelumnya yang akan dilupakan atau dihapus. *Forget gate* dapat dihitung menggunakan persamaan 1 di bawah ini.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

Keterangan:

f_t = Forget gate waktu ke-t

σ = Sigmoid (nilai dalam rentang 0-1)

w_f = Bobot

h_{t-1} = Hidden state sebelumnya

x_t = Input waktu ke-t

b_f = Bias

2. Input Gate

Berfungsi untuk mengontrol informasi baru dari input saat ini yang akan disimpan dalam memori. *Input gate* dapat dihitung menggunakan persamaan 2 di bawah ini.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

Analisis Sentimen Program Sekolah Rakyat Menggunakan Metode Bidirectional Long Short Term Memory

Keterangan:

i_t = *Input gate* waktu ke-t

σ = Sigmoid (nilai dalam rentang 0-1)

w_i = Bobot

h_{t-1} = *Hidden state* sebelumnya

x_t = *Input* ke-t

b_i = Bias

3. Candidate Cell State

Berfungsi untuk membentuk kandidat nilai baru yang akan ditambahkan ke memori. *Candidate cellsState* dapat dihitung menggunakan persamaan 3 di bawah ini.

$$\tilde{c}_t = \tanh(W_c \cdot [h_{t-1}, X_t] + b_c) \quad (3)$$

Keterangan:

\tilde{c}_t = Candidate cell state

\tanh = Fungsi aktivasi (rentang nilai -1 sampai 1)

w_c = Bobot

h_{t-1} = Hidden state sebelumnya

x_t = Input ke-t

b_c = Bias

4. Update Cell (Cell State)

Berfungsi untuk memperbarui memori dengan menggabungkan informasi dari memori sebelumnya dan kandidat baru. *Cell state* dapat dihitung menggunakan persamaan 4 di bawah ini.

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \quad (4)$$

Keterangan:

c_t = *Cell state*

f_t = *Forget gate* ke-t

c_{t-1} = *Cell state* sebelumnya

i_t = *Input gate* ke-t

\tilde{c}_t = *Candidate cell state* ke-t

5. Output Gate

Berfungsi untuk mengontrol informasi dari sel memori saat ini yang akan digunakan untuk menghitung *hidden state*. *Output gate* dapat dihitung menggunakan persamaan 5 di bawah ini.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

Keterangan:

o_t = *Output gate*

σ = Sigmoid (nilai dalam rentang 0-1)

w_o = Bobot

h_{t-1} = *Hidden state* sebelumnya

x_t = *Input* ke-t

b_o = Bobot

6. Hidden State

Memori jangka pendek dari data yang sedang diproses. *Hidden state* dapat dihitung menggunakan persamaan 6 di bawah ini.

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

Keterangan:

h_t = *Hidden state*

\tanh = Fungsi aktivasi (rentang nilai -1 sampai 1)

o_t = *Ouput gate* ke-t

C_t = *Cell state* ke-t

2.5 Pengumpulan Data

Data sentimen yang digunakan dalam penelitian ini merupakan data yang diambil dari platform X (x.com). Data ini didapatkan dalam rentang waktu 14 Juli 2025 sampai 20 September 2025 menggunakan metode *Crawling Data* dengan *tool tweet harvest* versi 2.6.1. Jumlah data yang didapatkan dari hasil *crawling data* pada platform X secara keseluruhan sebanyak 4714 baris data. Data tersebut kemudian dilakukan pelabelan secara manual oleh anotator. Setelah proses pelabelan data secara manual sudah selesai dilakukan kemudian masuk ke tahapan pembagian data, data-data tersebut nantinya akan dibagi menjadi tiga buah dataset yaitu ada: data pelatihan, data validasi, dan data pengujian dengan perbandingan 80:10:10.

2.6 Preprocessing

Text Preprocessing merupakan tahapan untuk memproses data tidak terstruktur menjadi data terstruktur (Wiujianna, Sunarno, & Iqbal, 2025). Berikut ini adalah tahapan dari text preprocessing:

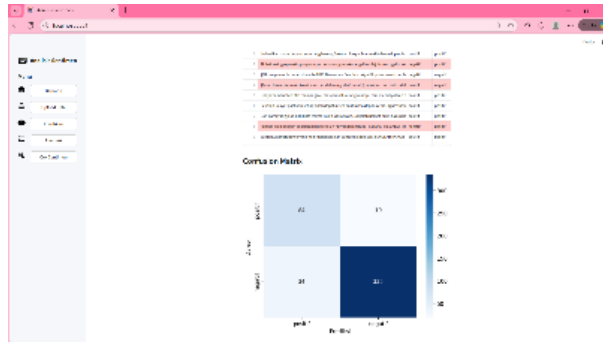
1. *Handling Missing Value*, proses ini digunakan untuk memastikan tidak ada nilai yang hilang dalam dataset. Dalam dataset yang digunakan proses ini dilakukan dengan cara menghapus baris data atau kolom data yang kosong dan tidak mengisi atau menggantinya dengan nilai apapun. Sedangkan dalam dataset yang berbeda bisa dilakukan dengan cara mengisi nilai kosong dengan nilai median, mean ataupun modus. (Kurniawan, 2024).
2. *Cek Duplikasi Data*, proses ini agar dapat dipastikan tidak ada lagi data atau baris ganda atau berulang pada data yang akan digunakan. Jika terdapat duplikasi data pada dataset maka duplikasi tersebut dapat dihapus (Zaza, Umar, & Sanga, 2024).
3. *Casefolding*, proses ini merupakan proses untuk menyamakan format teks pada dataset dengan mengubahnya menjadi huruf kecil (*lowercase*) (Pitriyanto, Ramadhan, Hakim, & Saputra, 2022).
4. *Data Cleaning*, proses ini dilakukan agar data yang akan digunakan atau diproses nantinya bersih dari elemen yang tidak diperlukan, contohnya seperti emotikon, url, mention atau tanda baca khusus lainnya. Elemen tersebut akan dihapus dari data karena tidak akan diproses. Dan jika tidak dihapus maka akan menimbulkan noise pada data yang akan mengganggu proses klasifikasi (Pamungkas & Februariyanti, 2022).
5. *Normalisasi*, proses ini dilakukan untuk menstandarkan teks pada data menjadi satu standar atau menyamakan kata pada teks atau data yang digunakan. Proses ini biasanya memanfaatkan kamus kata yang telah dikembangkan sebelumnya atau dapat juga membuat kamus sendiri (Yulianto, Hana, & Prihandono, 2025).
6. *Tokenisasi*, proses mengubah teks menjadi bagian kecil kata perkata yang bisa disebut dengan token (Angkoso, Thrisna, Satoto, & Kusumaningsih, 2024).
7. *Hapus Stopwords*, *stopwords* adalah kumpulan dari kata yang tidak membawa informasi atau makna yang penting sehingga dianggap tidak perlu diindex. Contoh seperti kata yang, itu, dan, dll. (Nurrochmah, Rahaningsih, Dana, & Rohmat, 2025).
8. *Stemming*, proses mengubah kata menjadi bentuk dasarnya (kata dasar) (Santosa, Purnamasari, & Mayasari, 2022).

2.7 Evaluasi

Pada proses evaluasi model menggunakan *confusion matrix*. *Confusion matrix* adalah matrik yang berisi tampilan dari perbandingan antara nilai sebenarnya atau label aktual dan nilai prediksi atau label prediksi dari model yang dikembangkan yang diperoleh selama proses testing berlangsung. *Confusion matrix* inilah yang nantinya digunakan untuk menghitung matrix performa seperti akurasi, presisi dan *recall* (Cahyana, Susanto, & Parjito, 2025). Terdapat 4 istilah dalam confusion matrix

Gambar 5 merupakan tampilan halaman klasifikasi. Pada halaman klasifikasi terjadi beberapa proses, pertama merupakan proses *preprocessing*, kemudian dilanjutkan dengan *label encoding*, setelah itu dilanjutkan dengan proses *balancing* data, kemudian dilanjutkan proses pelatihan dan pengujian. Jika hasil yang didapatkan kurang sesuai pengguna bisa mengatur parameter yang digunakan melalui menu pengaturan klasifikasi yang berada pada sidenar.

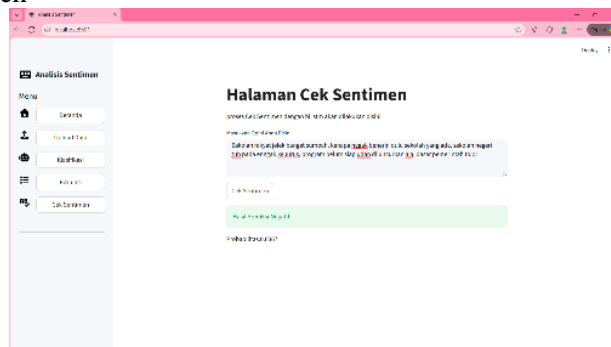
4. Halaman Evaluasi



Gambar 6. Halaman proses evaluasi

Gambar 6 menunjukkan hasil tampilan dari halaman evaluasi yang ada pada menu Evaluasi. Halaman evaluasi ini menampilkan hasil prediksi model terhadap data pengujian, dimana hasil prediksi yang tidak sesuai dengan label aktual akan dikenakan warna merah pada baris datanya. Kemudian ditampilkan visualisasi dari *confusion matrix* dan juga matrik performa model yang sudah dihitung berdasarkan pada *confusion matrix* tersebut.

5. Halaman Cek Sentimen



Gambar 7. Halaman cek sentimen

Gambar 7 merupakan tampilan dari halaman Cek Sentimen. Pada halaman ini pengguna bisa menginputkan opini mereka kemudian model akan mengklasifikasikan opini mereka untuk membuktikan keberhasilan model dalam melakukan klasifikasi.

3.2 Hasil Preprocessing

1. Casefolding

Tabel 1. Hasil casefolding

| Sebelum | Sesudah |
|-----------------------------------|-----------------------------------|
| @Airputih: | @airputih: |
| Malah bikin kesenjangan sosial 🚧🚧 | malah bikin kesenjangan sosial 🚧🚧 |

Pada tabel 1 terlihat bahwa setelah dilakukan proses casefolding semua karakter huruf menjadi lowercase.

2. Cleaning

Tabel 2. Hasil cleaning

| Sebelum | Sesudah |
|---------|---------|
|---------|---------|

Analisis Sentimen Program Sekolah Rakyat Menggunakan Metode Bidirectional Long Short Term Memory

@airputih: malah
malah bikin bikin kesenjangan
kesenjangan sosial
sosial 🤔 🤔

Pada tabel 2 terlihat sebelum dilakukan cleaning ,masih terdapat elemen mention dan emoticon pada teks. Setelah dilakukan proses cleaning teks sudah bersih dari elemen yang tidak perlukan yaitu mention (@) dan juga emoticon

3. Tokenisasi

Tabel 3. Hasil tokenisasi

| Sebelum | Sesudah |
|---|--|
| malah bikin kesenjangan sosial | malah,bikin, kesenjangan, sosial |

Pada tabel 3 terlihat setelah dilakukan tokenisasi teks akan dipecah menjadi token (kata per kata).

4. Normalisasi (Standarisasi Teks)

Tabel 4. Hasil normalisasi

| Sebelum | Sesudah |
|--|---|
| kenapa,nama,program,prabowo, itu,jadul,banget,koperasi,merah, putih,sekolah,rakyat,kek,ikutin, zaman,napa,sih | kenapa,nama,program,prabowo, itu,jadul,banget,koperasi,merah,putih, sekolah,rakyat,kaya,ikutin, zaman,kenapa,sih |

Pada tabel 4 terlihat setelah dilakukan proses normalisasi kata yang awalnya "napa" distandarkan atau dinormalisasikan menjadi "kenapa" menggunakan kamus normalisasi.

5. Hapus Stopwords

Tabel 5. Hasil hapus stopwords

| Sebelum | Sesudah |
|---|---|
| belajar,di, lingkungan, yang,suportif, bikin,sekolah, rakyat,semakin, keren,apalagi,bareng,prabowo | belajar,lingkungan,suportif, bikin,sekolah,rakyat, keren,bareng,prabowo |

Pada table 5 terlihat bahwa sebelum dilakukan proses hapus stopwords masih terdapat stopwords yaitu kata: "di", "yang", "semakin". Seteklah dilakukan proses hapus stopwords maka stopwords yang ada sudah dihapus.

6. Stemming

Tabel 6. Hasil stemming

| Sebelum | Sesudah |
|--------------------------|----------------------|
| bikin,kesenjangan,sosial | bikin,senjang,sosial |

Pada tabel 6 terlihat bahwa sebelum dilakukan proses *stemming* kata "kesenjangan" merupakan kata yang berimbuhan ke dan an. Setelah dilakukan proses *stemming* kata tersebut diubah ke bentuk dasarnya menjadi senjang.

3.3 Label Encoding

Tabel 7. Hasil label encoding

| final_text | aktual label | Label encoded |
|------------|--------------|---------------|
| | | |

| | |
|---|---|
| mantep banget positif | 1 |
| nih program sekolah rakyat prabowo anak anak keluarga didik kualitas pimpin peduli bangsa | |
| bangun sekolah negatif | 0 |
| rakyat yaudah dipaksain fokus maksimalin mah merata timpang aduh marah banget banget banget | |

3.4 Evaluasi

Tabel 8. Confusion matrix

| Aktual | Predict | | Total |
|---------|----------|---------|-------|
| | Positif | Negatif | |
| Positif | 335 (TP) | 34 (FN) | 369 |
| Negatif | 19 (FP) | 84 (TN) | 103 |
| Total | 354 | 118 | 472 |

Tabel 8 merupakan *confusion matrix* dari data pelatihan. Total data pelatihan sebanyak 472 dengan distribusi 369 positif dan 103 negatif. *Confusion matrix* inilah yang kemudian akan digunakan menjadi bahan acuan atau referensi untuk melakukan perhitungan matrik performa model di bawah ini.

1. Kelas Positif

$$\text{Precision} = \frac{TP}{TP+FP} = \frac{335}{335+19} = \frac{335}{354} = 0.9463 \text{ (7)}$$

$$\text{Recall} = \frac{TP}{TP+FN} = \frac{335}{335+34} = \frac{335}{369} = 0.908 \text{ (8)}$$

Hasil dari perhitungan matrik performa untuk kelas positif mencapai precision 0.9463 ditunjukkan pada persamaan 7 dengan *recall* mencapai 0.908 ditunjukkan pada persamaan 8, angka tersebut membuktikan bahwa model dapat mengenali data positif dengan baik karena tingginya nilai presisi dan *recall*.

2. Kelas Negatif

$$\text{Precision} = \frac{TN}{TN+FN} = \frac{84}{84+34} = \frac{84}{118} = 0.7119 \text{ (9)}$$

$$\text{Recall} = \frac{TN}{TN+FP} = \frac{84}{84+19} = \frac{84}{103} = 0.8155 \text{ (10)}$$

Hasil perhitungan matrik performa untuk kelas negatif mencapai precision 0.7119 ditunjukkan pada persamaan 9 dengan *recall* mencapai 0.8115 ditunjukkan pada persamaan 10. Hasil perhitungan dari presisi dan recall yang didapat pada kelas negatif jauh lebih rendah jika dibandingkan dengan hasil perhitungannya presisi dan *recall* yang didapatkan dari kelas positif, bahkan jika dihitung jaraknya terpaut hampir mencapai 20%. Tentu saja hal tersebut amat sangat dimungkinkan jika dilihat dari distribusi data kelas positif dan negatif yang tidaklah sama dan bahkan terpaut jauh. Namun, meski demikian model tetap dapat mengklasifikasikan data dengan baik, dilihat juga dari hasil *confusion matrix* data negatif yang berhasil diprediksi benar sebanyak 84 data dari 103 data.

3. Weighted

$$\begin{aligned} \text{Precision} &= \frac{(369 \times 0.9463) + (103 \times 0.7119)}{472} \\ &= \frac{349.2 + 73.3}{472} = 0.895 \quad (11) \end{aligned}$$

$$\begin{aligned} \text{Recall} &= \frac{(369 \times 0.908) + (103 \times 0.8155)}{472} \\ &= \frac{335 + 84}{472} = 0.887 \quad (12) \end{aligned}$$

Weighted merupakan hasil precision dan recall dari keseluruhan kelas. Hasil dari perhitungan kelas secara keseluruhan (positif dan negatif) membuktikan bahwa model memperoleh presisi sebesar 0.895 ditunjukkan pada persamaan 11 dengan *recall* sebesar 0.887 ditunjukkan pada persamaan 12. Hal tersebut membuktikan bahwa kinerja model yang sangat efektif dalam mengenali kelas positif dan negatif serta melakukan klasifikasi

4. Akurasi

$$\frac{TP+TN}{TP+TN+FP+FN} = \frac{335+84}{335+84+19+34} = \frac{419}{472} = 0.8881 \quad (13)$$

Hasil perhitungan akurasi mencapai 0.8881 ditunjukkan pada persamaan 13. Hal ini menunjukkan bahwa model dapat mengenali semua kelas (positif dan negatif) hampir 89%. Sehingga dapat dipastikan model dapat melakukan klasifikasi dengan baik.

5. Classification Report

Tabel 9. Classification report

| Metrik | Positif | Negatif | Weighted |
|-----------|---------|---------|----------|
| Precision | 0.9463 | 0.7119 | 0.895 |
| Recall | 0.908 | 0.8155 | 0.887 |
| Accuracy | | | 0.8881 |

Tabel 9 merupakan hasil keseluruhan evaluasi pada masing-masing kelas dan juga keseluruhan kelas. Tabel ini menampilkan hasil akurasi yang sudah dicapai model, kemudian hasil perhitungan matrik performa yaitu presisi dan *recall* pada masing-masing kelas, dan juga keseluruhan kelas.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan mayoritas sentimen masyarakat terkait program sekolah rakyat adalah positif dengan jumlah 3683. Hal ini membuktikan bahwa tanggapan atau opini masyarakat mengenai adanya program sekolah rakyat yang sudah belangsung saat ini merupakan tanggapan yang mendukung keberlangsungan program sekolah rakyat kedepannya atau berupa afirmasi positif. Sedangkan sentimen negatif berjumlah 1031, diantaranya berisi tentang sistem pendidikan yang harus dibenahi, pemilihan lokasi sekolah rakyat, dan proses masuk peserta didik. Hal lain yang berkaitan juga sudah semestinya menjadi perhatian masyarakat bersama terutama pihak-pihak berwenang yang terkait dalam mewujudkan pendidikan yang setara dan merata di seluruh Indonesia, karena pendidikan sendiri adalah hak segala bangsa tanpa terkecuali. Model BiLSTM juga dapat mengklasifikasikan sentimen dengan baik dilihat dari nilai matrik performa yang dihasilkan, nilai akurasi, presisi dan *recall* mencapai 89%. Meskipun jika dilihat per kelas nilai matrik performa kelas negatif lebih rendah dikarenakan perbedaan jumlah data.

5. SARAN

Harapannya kedepan adalagi menelitian terkait program ini yang dimana datanya tidak hanya diambil dari satu platform saja melainkan dari beberapa platform agar data yang didapatkan lebih kompleks karena banyak buzzer media sosial yang dibayar untuk membuat opini.

DAFTAR PUSTAKA

- [1] B. Setpres, "Prabowo Subianto dan Gibran Rakabuming Raka Resmi Dilantik Sebagai Presiden dan Wakil Presiden RI," 20 Oktober 2024. [Online]. Available: <https://www.presidentri.go.id/siaran-pers/prabowo-subianto-dan-gibran-rakabuming-raka-resmi-dilantik-sebagai-presiden-dan-wakil-presiden-ri/>.
- [2] S. Noerdin, "Sekolah Rakyat Indonesia: Terobosan Pendidikan Gratis atau Kebijakan Kontroversial?," 2025. [Online]. Available: <https://umj.ac.id/opini/sekolah-rakyat-indonesia-terobosan-pendidikan-gratis-atau-kebijakan-kontroversial/>.
- [3] D. T. Attaulah and D. Soyusiawaty, "Analisis Sentimen Program Makan Siang Gratis di Twitter/X menggunakan Metode BI-LSTM," *Edumatic: Jurnal Pendidikan Informatika*, vol. XI, no. 01, pp. 294-303, 2025.
- [4] F. Bouchra, I. M. A. D. Suarjaya and N. K. D. Rusjyanthi, "Analisis Sentimen Masyarakat terhadap Tayangan Televisi Nasional menggunakan Metode Deep Learning," *Jurnal Buana Informatika*, pp. 89-99, 2024.
- [5] A. Wiujianna, Sunarno and Iqbal, "Perbandingan Performa Model Long Short-Term Memory dan Bidirectional untuk Prediksi Kabut," *Jurnal Teknik Informatika dan Sistem Informasi (JuTISI)*, vol. XI, no. 02, pp. 251-260, 2025.
- [6] S. Y. Kurniawan, "Klasifikasi Kelayakan Air Minum Dengan Backpropagation Neural Network Berbasis Penanganan Missing Value Dan Normalisasi," *Journal of Information System Research (JOSH)*, vol. VI, no. 01, pp. 87-95, 2024.
- [7] D. A. U. Zaza, E. Umar and F. E. O. Sanga, "Penerapan Text Mining Dalam Klasifikasi Judul Skripsi Mahasiswa Pada Universitas Stella Maris Sumba Menggunakan Metode Naïve Bayes," *Journal Of Informatics And Business*, vol. II, no. 03, pp. 327-337, 2024.
- [8] A. Pitriyanto, M. N. Ramadhan, T. A. Hakim and A. Saputra, "Analisis Sentimen Pembelajaran Daring Pada Masa Pandemi Covid-19 di Indonesia," *Bangkit Indonesia*, vol. XI, no. 01, pp. 7-13, 2022.
- [9] M. D. Pamungkas and H. Februariyanti, "Penerapan Algoritma K-Means Clustering Untuk Mengelompokkan Data Review Barang Pada E-Commerce Lazada," *semanTIK*, vol. VIII, no. 02, pp. 99-104, 2022.
- [10] M. F. Yulianto, F. M. Hana and A. Prihandono, "Penerapan Algoritma Naïve Bayes Dalam Analisis Sentimen terhadap Mobil Listrik," *Sainteks*, vol. XXII, no. 01, pp. 109-115, 2025.
- [11] C. V. Angkoso, M. A. N. Thrisna, B. D. Satoto and A. Kusumaningsih, "Optimasi Klasifikasi Sentimen Menggunakan Random Forest dengan Preprocessing K-Means Clustering dan SMOTE," *JEPIN (Jurnal Edukasi dan Penelitian Informatika)*, vol. X, no. 03, pp. 389-400, 2024.
- [12] D. S. Nurrochmah, N. Rahaningsih, R. D. Dana and C. L. Rohmat, "Penerapan Algoritma Naive Bayes Dalam Analisis Sentimen Ulasan Aplikasi Kitalulus Di Google Play Store," *Jurnal InformatikaTerpadu*, vol. X, no. 03, pp. 1-11, 2025.
- [13] A. Santosa, I. Purnamasari and R. Mayasari, "Pengaruh Stopword Removal dan Stemming Terhadap Performa Klasifikasi Teks Komentar Kebijakan New Normal Menggunakan Algoritma LSTM," *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. VI, no. 01, pp. 81-93, 2022.
- [14] A. Cahyana, E. R. Susanto and Parjito, "Penerapan Algoritma XGBoost untuk Prediksi Diabetes: Analisis Confusion Matrix dan ROC Curve," *Fountain of Informatics Journal*, pp. 40-50, 2025.
- [15] M. K. Suryadewiansyah and T. E. E. Tju, "Naïve Bayes dan Confusion Matrix untuk Efisiensi Analisa Intrusion Detection System Alert," *Jurnal Nasional Teknologi dan Sistem Informasi*, pp. 81-88, 2022.