

Penerapan *Long Short-Term Memory* dan TF-IDF dalam Mengklasifikasikan Sentimen Publik terhadap Kebijakan Efisiensi Anggaran 2025

Hamidah Lutfiyanti Maharani ^{#1}, Mokhammad Amin Hariyadi ^{#2}, Zainal Abidin^{#3}

^{#1,2,3}Teknik Informatika, UIN Maulana Malik Ibrahim Malang, Malang, Indonesia

Korespondensi author ^{*}adyt2002@uin-malang.ac.id

Info Artikel

Diajukan: 28 Juli 2025

Diterima: 22 Januari 2026

Diterbitkan: 27 Januari 2026

Keywords:

Budget Efficiency 2025, Sentiment Classification, LSTM, TF-IDF

Kata Kunci:

Efisiensi Anggaran 2025, Klasifikasi Sentimen, LSTM, TF-IDF



Lisensi: cc-by-sa

Copyright © 2025 Hamidah Lutfiyanti Maharani, Mokhammad Amin Hariyadi, Zainal Abidin

Abstract

The 2025 Budget Efficiency Policy has generated diverse public responses, widely expressed through social media, making sentiment analysis necessary to understand overall public opinion. The main problem addressed in this study is how to accurately classify public sentiment toward the policy using unstructured text data. This research aims to classify public sentiment into positive and negative categories and to evaluate the performance of the Long Short-Term Memory (LSTM) model combined with Term Frequency-Inverse Document Frequency (TF-IDF). The research method involves collecting 2,222 tweets from the X platform, manual data labeling, text preprocessing, TF-IDF weighting, and sentiment classification using the LSTM model. Model performance is evaluated using accuracy, precision, recall, and F1-score across several data-splitting scenarios. The results show that the combination of LSTM and TF-IDF performs effectively in sentiment classification. The best performance is achieved with an 80:20 data split, 10 epochs, and a batch size of 32, resulting in an accuracy of 94.38%, precision of 93.75%, recall of 89.44%, and F1-score of 91.55%. These findings indicate that LSTM is well suited for analyzing public sentiment toward government policies based on social media data.

Abstrak

Kebijakan Efisiensi Anggaran 2025 memunculkan beragam respons publik yang tersebar luas di media sosial, sehingga diperlukan analisis sentimen untuk memahami kecenderungan opini masyarakat. Permasalahan dalam penelitian ini adalah bagaimana mengklasifikasikan sentimen publik terhadap kebijakan tersebut secara akurat menggunakan data teks yang tidak terstruktur. Penelitian ini bertujuan mengklasifikasikan sentimen publik ke dalam kategori positif dan negatif serta mengevaluasi kinerja model Long Short-Term Memory (LSTM) dengan pembobotan Term Frequency-Inverse Document Frequency (TF-IDF). Metode penelitian meliputi pengumpulan 2.222 tweet dari platform X, pelabelan data secara manual, preprocessing teks, pembobotan kata menggunakan TF-IDF, serta klasifikasi sentimen menggunakan model LSTM. Evaluasi dilakukan dengan metrik akurasi, presisi, recall, dan F1-score pada beberapa skenario pembagian data. Hasil penelitian menunjukkan bahwa kombinasi LSTM dan TF-IDF mampu mengklasifikasikan sentimen publik dengan baik. Kinerja terbaik diperoleh pada rasio data 80:20 dengan epoch 10 dan batch size 32, menghasilkan akurasi 94,38%, presisi 93,75%, recall 89,44%, dan F1-score 91,55%. Temuan ini membuktikan efektivitas LSTM dalam analisis sentimen kebijakan publik berbasis media sosial.

Cara mensitasi artikel:

H. L. Maharani, M. A. Hariyadi, Z. Abidin. "Penerapan Long Short-Term Memory dan TF-IDF dalam Mengklasifikasikan Sentimen Publik terhadap Kebijakan Efisiensi Anggaran 2025." Jurnal Teknologi Informasi: Teori, Konsep, dan Implementasi (JTI-TKI), vol. 16, no. 2, pp. 77-85, Oktober 2025, <https://doi.org/10.36382/jti-tki.v16i2.605>

PENDAHULUAN

Pengelolaan anggaran negara yang efisien sangat penting untuk keberlanjutan pembangunan suatu negara [1]. Efisiensi anggaran bertujuan untuk memaksimalkan penggunaan dana negara dengan cara mengurangi pemborosan dan memastikan setiap rupiah yang dibelanjakan memberikan manfaat maksimal bagi masyarakat. Efisiensi ini tidak hanya mencakup pemangkasan anggaran, tetapi juga memastikan bahwa dana yang tersedia dialokasikan dengan tepat untuk sektor-sektor yang memiliki dampak besar terhadap kesejahteraan publik [2]. Dengan pengelolaan yang baik, setiap alokasi anggaran dapat memberikan hasil yang optimal dalam mendukung pembangunan berkelanjutan. Di tengah era digital saat ini, di mana opini publik dengan

cepat tersebar di media sosial, memahami respons masyarakat terhadap kebijakan tersebut menjadi semakin penting.

Pada pemerintahan era baru, terdapat kebijakan efisiensi anggaran 2025 yang dirancang untuk mengatasi tantangan fiskal yang dihadapi negara, dengan target penghematan sebesar Rp306,69 triliun [3]. Kebijakan ini mencakup pemangkasan anggaran untuk kementerian/lembaga serta pengalihan sebagian dana ke daerah, dengan fokus utama pada program-program prioritas. Tujuan dari kebijakan ini adalah untuk memastikan bahwa anggaran negara digunakan lebih efektif dan tepat sasaran, mendukung pembangunan yang lebih berkelanjutan. Namun, meskipun bertujuan untuk memperbaiki pengelolaan keuangan negara, kebijakan ini menghadirkan potensi dampak yang besar, terutama bagi

sektor-sektor yang sangat bergantung pada dana publik, seperti pendidikan dan kesehatan [4]. Hal ini menimbulkan kekhawatiran di kalangan masyarakat yang merasakan dampak langsung, baik melalui pemangkasan program sosial maupun berkurangnya alokasi anggaran untuk layanan publik.

Sebagian masyarakat mendukung kebijakan ini karena diyakini dapat menjaga kestabilan fiskal negara dan mengarahkan dana ke sektor yang lebih penting. Di sisi lain, banyak pula yang merasa khawatir dengan pemangkasan anggaran, terutama yang berdampak pada sektor sosial dan publik. Pemangkasan anggaran untuk program sosial, seperti bantuan langsung dan pendidikan, dapat memengaruhi kualitas layanan yang diterima oleh masyarakat [5]. Kekhawatiran ini muncul di tengah meningkatnya ketergantungan masyarakat terhadap bantuan sosial dan kualitas layanan publik yang diharapkan tetapi stabil meski ada penghematan anggaran. Hal ini mengundang pro dan kontra di kalangan masyarakat, dengan beberapa pihak merasa terancam oleh kebijakan ini. Sebagian besar kritikan muncul dari sektor yang merasa langsung terdampak, seperti sektor kesehatan dan pendidikan. Oleh karena itu, penting untuk mengetahui bagaimana masyarakat memandang kebijakan ini melalui tanggapan yang mereka sampaikan di media sosial.

Media sosial, khususnya X, menjadi platform utama dengan 99,55% dari total sosial media yang ada untuk menyampaikan pendapat mereka mengenai kebijakan pemerintah [6]. X memungkinkan masyarakat untuk berbagi pandangan mereka secara langsung dan cepat mengenai berbagai isu, termasuk kebijakan efisiensi anggaran 2025. Dalam konteks ini, media sosial bukan hanya menjadi sarana untuk berbagi informasi, tetapi juga sebagai indikator cepat untuk menilai respons publik terhadap kebijakan-kebijakan pemerintah. Dengan adanya platform ini, pemerintah dapat memperoleh gambaran lebih jelas mengenai pandangan publik terhadap kebijakan yang diterapkan. Tanggapan yang diberikan masyarakat di X mencakup berbagai opini yang mencerminkan sentimen positif dan negatif terkait kebijakan ini [7]. Namun, karena data yang terkumpul bersifat tidak terstruktur, diperlukan metode yang tepat untuk mengklasifikasikan sentimen yang terkandung dalam *tweet* tersebut. Salah satu metode yang dapat digunakan untuk analisis ini adalah model *Long Short-Term Memory* (LSTM).

Long Short-Term Memory (LSTM) adalah jenis jaringan saraf tiruan yang dirancang untuk memproses data sekuensial, seperti teks [8]. Model ini memiliki kemampuan untuk mengatasi masalah vanishing gradient yang sering terjadi pada *Recurrent Neural Networks* (RNN), sehingga sangat efektif untuk menganalisis teks yang bersifat panjang dan kompleks [9]. LSTM dapat menangkap informasi jangka panjang dalam urutan data, yang memungkinkan model untuk memahami konteks yang lebih dalam dari setiap sentimen yang diekspresikan. Metode ini telah terbukti efektif dalam berbagai penelitian sebelumnya yang mengkaji sentimen publik terhadap kebijakan pemerintah. Misalnya, penelitian yang

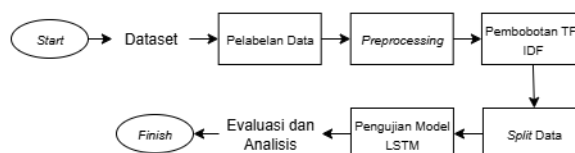
dilakukan pada tahun 2021 [10] menggunakan algoritma *Naïve Bayes* untuk menganalisis kebijakan kartu prakerja dan mencapai akurasi sebesar 91,06%. Penelitian lainnya oleh Mahmuji dkk juga menggunakan LSTM untuk menganalisis sentimen terhadap penundaan Pemilu 2024, yang menunjukkan keunggulan LSTM dengan akurasi 92% [11]. Dari penelitian tersebut terbukti LSTM lebih baik dalam mengklasifikasikan sentimen dalam bentuk teks.

Dalam penelitian ini, data yang digunakan adalah *tweet* yang terkait dengan kebijakan efisiensi anggaran 2025. Data ini kemudian akan melalui tahap *preprocessing*, yang meliputi pembersihan data (*data cleaning*), pengubahan teks menjadi format yang konsisten (*case folding*), normalisasi kata, penghilangan kata yang tidak relevan (*stopword removal*), dan pengubahan kata ke bentuk dasar (*stemming*) [12]. Setiap tahap *preprocessing* bertujuan untuk memastikan bahwa data yang digunakan dalam analisis bersih dari elemen yang tidak perlu dan siap untuk diproses lebih lanjut. Setelah data diproses, tahap selanjutnya adalah pelatihan model LSTM dengan data yang telah dilabeli dengan kategori sentimen positif dan negatif. Model yang telah dilatih akan digunakan untuk mengklasifikasikan sentimen dalam *tweet* masyarakat terkait kebijakan ini.

Dengan melalui beberapa proses *preprocessing* serta penggunaan model *Long Short-Term Memory*, penelitian ini mengklasifikasi sentimen untuk menggambarkan dukungan atau penolakan terhadap kebijakan pemerintahan tersebut. Selain itu, penelitian ini juga mengidentifikasi topik-topik utama yang menjadi perhatian masyarakat terkait kebijakan ini. Hasil penelitian diharapkan memberikan wawasan bagi pemerintah untuk mengevaluasi dan merumuskan kebijakan yang lebih tepat di masa depan.

METODE

Penelitian ini melalui beberapa tahapan, mulai dari pengumpulan data, *preprocessing*, hingga pelatihan dan evaluasi model. Setiap tahapan dilakukan untuk memastikan data yang digunakan bersih dan siap untuk analisis. Alur dari penelitian ini dapat dilihat pada **Gambar 1** berikut.



Gambar 1. Alur Penelitian

Berdasarkan **Gambar 1**, penelitian ini dimulai dengan pengumpulan dataset berupa *tweet* yang digunakan sebagai input analisis. Dikarenakan *tweet* tersebut belum dilabeli, langkah pertama adalah pelabelan data secara manual. Setelah pelabelan, data kemudian diproses melalui tahap *preprocessing* untuk mengurangi *noise* dan

mempersiapkan data agar siap untuk analisis lebih lanjut. Data yang telah diproses selanjutnya diberi pembobotan menggunakan metode TF-IDF untuk mengubah teks menjadi format numerik. Kemudian, data dibagi menjadi dua bagian: data pelatihan dan data pengujian. Data pelatihan digunakan untuk melatih model LSTM, yang selanjutnya dievaluasi untuk mengukur kinerja dan efektivitas model dalam mengklasifikasikan sentimen.

A. Dataset

Dataset yang digunakan tersedia dalam dataset berjudul dataset yang tersedia di Kaggle dengan judul "*Budget Efficiency in Indonesia*" yang dapat diakses melalui platform Kaggle dengan tautan sebagai berikut: <https://www.kaggle.com/datasets/jocelyndumlao/budget-efficiency-in-indonesia>. Pemilik dari data tersebut adalah Asno Azzawagama Firdaus [13] dengan cara pengambilan data *crawling* dari media sosial X. Sehingga *tweet* yang ada pada X sesuai dengan kriteria akan dikumpulkan. Keseluruhan *tweet* mencakup pembahasan terkait dengan kebijakan Efisiensi Anggaran 2025, yang diunggah oleh masyarakat sebagai respons terhadap kebijakan tersebut.

Jumlah total *tweet* yang terdapat dalam dataset ini adalah 2.222 *tweet*, yang mencerminkan berbagai pandangan dan tanggapan publik terkait kebijakan Efisiensi Anggaran 2025. Data ini bersifat tidak terstruktur dan berisi teks mentah yang mencakup alasan masyarakat mendukung atau menentang kebijakan, serta keluhan mengenai dampaknya. Mayoritas *tweet* mengungkapkan kekhawatiran terhadap pemangkasan anggaran yang memengaruhi sektor vital, seperti pendidikan dan kesehatan, sementara sebagian lainnya mendukung kebijakan efisiensi untuk menjaga kestabilan fiskal negara.

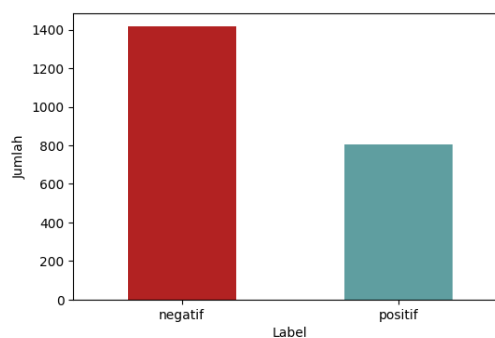
B. Pelabelan Data

Setelah data terkumpul, langkah selanjutnya adalah pelabelan data. Label yang digunakan yaitu positif dan negatif. Pelabelan dilakukan secara manual oleh penulis untuk setiap *tweet* yang ada, kemudian hasil pelabelan tersebut divalidasi oleh seorang *validator* yang ahli dalam bahasa. Proses ini memastikan bahwa data yang dilabeli sesuai dengan kaidah bahasa yang tepat dan dikategorikan dengan benar. Tabel 1 akan menampilkan beberapa contoh data yang digunakan pada penelitian.

Tabel 1 memperlihatkan data *tweet* serta label sesuai dengan kategorinya. Label positif menunjukkan *tweet* yang mendukung kebijakan Efisiensi Anggaran 2025, biasanya mencerminkan pandangan yang menganggap kebijakan tersebut sebagai langkah yang baik untuk kestabilan fiskal negara. Sementara itu, label negatif mengindikasikan *tweet* yang menentang kebijakan tersebut, dengan keluhan atau kekhawatiran terkait dampak pemangkasan anggaran terhadap sektor-sektor vital seperti pendidikan dan kesehatan. Berikut akan ditampilkan grafik dari setiap label pada **Gambar 2**.

Tabel. 1 Contoh Data

<i>Tweet</i>	Label
Kalo mau efisiensi anggaran tu ya wamen satu aja sama staf utusan khusus setara menteri itu dibubarin	Negatif
Bersama membangun kesejahteraan dengan efisiensi anggaran #AnggaranBijakUntukRakyat	Positif
@blackjasmineee @worksfess Btw anggaran efisiensi itu ga sama dengan anggaran CPNS ya. Selama belum ada info dari menpanRB/ pres kurang kurangilah fear mongering kayak gini.	Positif
@spectatorindex ALAMAK Audit Efisiensi Anggaran Trump luar biasa ngaco	Negatif



Gambar. 2 Grafik Jumlah Setiap Label

Dari **Gambar 2** grafik menunjukkan jumlah *tweet* negatif yang unggul mencapai nilai 1416 atau 63,73% dari total keseluruhan. Sedangkan *tweet* positif sebanyak 806 atau 36,27%.

C. Preprocessing

Preprocessing merupakan tahap yang penting dalam penelitian. Tujuannya untuk mempersiapkan data teks agar dapat digunakan secara optimal oleh model analisis sentimen. Dataset yang digunakan berasal dari *tweet-tweet* masyarakat terkait kebijakan Efisiensi Anggaran 2025. Data mentah yang dikumpulkan melalui media sosial X memerlukan serangkaian proses untuk membersihkan dan merapikan teks agar siap dianalisis lebih lanjut. Berikut adalah tahapan-tahapan preprocessing:

- *Data Cleaning*

Pada tahap ini, elemen-elemen yang tidak relevan seperti URL, angka, tanda baca, dan karakter khusus dihapus. Misalnya, *tweet* seperti "Kebijakan anggaran 2025 sangat bagus! <https://link.com>" akan dibersihkan menjadi "Kebijakan anggaran 2025 sangat bagus".

- *Case Folding*

Seluruh teks diubah menjadi huruf kecil untuk memastikan konsistensi dan mencegah perbedaan antara kata yang memiliki kapitalisasi berbeda. Sebagai contoh, "Anggaran" dan "anggaran" dianggap sama setelah proses ini.

- *Stopwords Removal*

Kata-kata umum yang tidak berkontribusi signifikan terhadap analisis, seperti "yang", "di", atau "untuk", dihapus. Sebagai contoh, kalimat "Kebijakan ini sangat bermanfaat untuk masyarakat" menjadi "Kebijakan bermanfaat masyarakat".

- *Normalisasi*

Pada tahap ini, kata-kata yang tidak baku atau bentuk singkatan diubah menjadi bentuk standar. Misalnya, kata "ttp" yang berarti "tetapi" akan digantikan dengan kata "tetapi" yang benar.

- *Stemming*

Stemming digunakan untuk mengubah kata-kata yang memiliki variasi bentuk ke bentuk dasarnya. Sebagai contoh, kata "memperbaiki" diubah menjadi "baik" untuk menyederhanakan analisis.

- *Tokenizing*

Pada tahap tokenisasi, teks dibagi menjadi unit-unit kecil yang disebut token, biasanya berupa kata. Misalnya, kalimat "Anggaran 2025 dipotong" akan dipisahkan menjadi "Anggaran", "2025", "dipotong".

Setelah melalui tahapan-tahapan tersebut, data yang sudah dibersihkan dan disiapkan akan dilanjutkan ke tahap selanjutnya yaitu pembobotan kata TF-IDF.

D. Pembobotan TF-IDF

Pembobotan data bertujuan untuk mengubah data teks menjadi bentuk numerik dengan menggunakan metode TF-IDF. Metode ini mengukur pentingnya sebuah kata dalam suatu dokumen dengan mempertimbangkan frekuensinya di dokumen tersebut [14]. Tidak hanya itu, dihitung juga seberapa jarang kata tersebut muncul diseluruh dataset, yang mencerminkan sejauh mana kata itu unik dan relevan. Dengan demikian, TF-IDF membantu model untuk fokus pada kata-kata yang lebih bermakna dalam analisis. Pembobotan ini terdiri dari dua komponen utama TF dan IDF (*Inverse Document Frequency*).

TF (*Term Frequency*) sebagai nilai untuk mengukur seberapa sering suatu kata muncul dalam sebuah dokumen. Semakin sering kata tersebut muncul, semakin tinggi nilai TF-nya. Rumus untuk menghitung TF adalah:

$$TF = \frac{\text{jumlah kemunculan kata dalam dokumen}}{\text{jumlah kata pada dokumen}}$$

IDF (*Inverse Document Frequency*) digunakan untuk mengukur seberapa penting atau unik suatu kata di seluruh koleksi dokumen. Kata yang muncul di lebih banyak dokumen memiliki nilai IDF yang rendah, sementara kata yang hanya muncul di sedikit dokumen memiliki nilai IDF yang lebih tinggi. Rumus untuk menghitung IDF adalah:

$$IDF = \log \frac{|D|}{DF}$$

Keterangan:

IDF : Frekuensi kemunculan kata terbalik

D : Jumlah total dokumen

DF : Jumlah dokumen yang mengandung kata

Setelah didapatkan nilai TF dan IDF, selanjutnya menggabungkan kedua nilai tersebut dengan mengkalikan agar mendapatkan nilai TF-IDF. Nantinya hasil dari perhitungan tersebut merupakan hasil numerik dari sebuah kata. Rumus dari TF-IDF sebagai berikut:

$$TF - IDF = TF \times IDF$$

Sebagai contoh, akan ditampilkan pada **Tabel 2** dari hasil perhitungan TF-IDF.

Tabel. 2 Contoh Pembobotan TF-IDF

Kata	D1	D2	D3
dampak	0	0	0
efisiensi	0	0	0
anggar	0	0	0
rakyat	0.0687	0	0.0126
butuh	0.0523	0	0
cari	0.0687	0	0
biaya	0.0523	0	0
hidup	0.0523	0.1570	0

Dari **Tabel 2** setiap kata memiliki nilai tersendiri setelah dilakukan pembobotan kata. Nilai 0 menunjukkan kata tersebut tidak terdapat dalam 1 dokumen.

E. Split Data

Dari keseluruhan data akan dilakukan pembagian data menjadi dua bagian yaitu *data training* (data pelatihan) dan *data testing* (data pelatihan). *Data training* digunakan untuk melatih model, dimana data ini akan membantu model mempelajari pola-pola dan hubungan yang ada di dalam dataset. Sedangkan *data testing* digunakan untuk menguji kemampuan model dalam menggeneralisasi dan memprediksi data yang belum pernah dilihat sebelumnya. Pembagian data untuk memastikan bahwa model tidak hanya mengingat data, tetapi juga dapat bekerja dengan baik pada data baru yang belum dipelajari. Dalam penelitian ini pembagian data akan dilakukan dalam 3 rasio yang akan ditampilkan pada **Tabel 3**.

Tabel. 3 Split Data

Rasio Data	Data Training	Data Testing
70 : 30	1555	667
80 : 20	1778	444
90 : 10	1999	223

Untuk membagi data, digunakan *random seed* agar pembagian data antara *training* dan *testing* dapat dilakukan secara acak namun konsisten. Penggunaan *random seed* memastikan bahwa pembagian data yang sama dapat diulang pada percobaan berikutnya, memberikan hasil yang dapat diuji secara reproduktif.

F. Model Long Short-Term Memory

Dalam pengklasifikasian sentimen digunakan model LSTM. Model ini merupakan jenis jaringan syaraf tiruan yang dirancang untuk memproses data sekuensial, seperti teks, dengan kemampuan mengingat informasi jangka

panjang. LSTM dilengkapi dengan gate. Terdapat tiga gate utama yang digunakan model LSTM yaitu [15]:

- *Forget Gate*

Gerbang ini menentukan informasi mana yang perlu dilupakan atau disimpan dari langkah sebelumnya. Informasi yang dianggap tidak relevan akan dihapus. Rumus yang digunakan yaitu:

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

- *Input Gate*

Setelah itu, ditentukan juga informasi baru yang akan disimpan dalam memori model. Ini berfungsi untuk memperbarui informasi berdasarkan input yang baru saja diproses. Rumus perhitungannya yaitu:

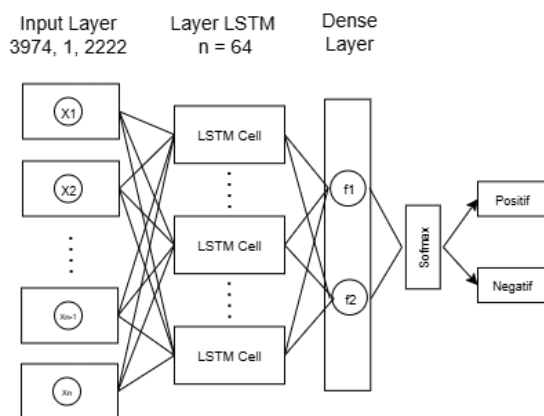
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

- *Output Gate*

Saat informasi sudah melewati semua gate, maka masuk ke dalam Output gate sebagai gerbang akhir untuk menentukan informasi mana yang akan diteruskan ke langkah berikutnya dalam jaringan. Output yang dihasilkan adalah informasi yang relevan berdasarkan semua perhitungan yang telah dilakukan sebelumnya. Berikut merupakan rumus perhitungannya:

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

Dalam memproses data digunakan arsitektur model LSTM sebagai struktur untuk membentuk jaringan model. Pada penelitian ini digunakan arsitektur model yang akan digambarkan pada **Gambar 3**.



Gambar. 3 Arsitektur LSTM

Gambar 3 menggambarkan arsitektur dari model LSTM yang digunakan dalam penelitian ini untuk mengklasifikasi sentimen. Setiap X_1, X_2, \dots, X_n mewakili input sekuensial, seperti kata-kata dalam sebuah tweet, yang dimasukkan ke dalam LSTM Cells satu per satu. Setiap LSTM Cell akan memproses input dan menghasilkan output berdasarkan informasi yang telah diproses sebelumnya, dengan mempertimbangkan dependensi jangka panjang dalam data. Setelah itu, hasil

dari seluruh LSTM Cells diteruskan ke *fully connected layer* yang menghasilkan dua output, yaitu f_1 dan f_2 , yang kemudian diproses menggunakan fungsi *Softmax* untuk mengklasifikasi sentimen sebagai Positif atau Negatif. Struktur ini memungkinkan model untuk memproses data teks secara efektif dengan mempertimbangkan konteks dari kata-kata yang ada dalam urutan tweet.

Dalam proses pelatihan model LSTM terdapat *hyperparameter* untuk mengendalikan serta mengatur bagaimana model dilatih dan dioptimalkan agar model tidak terjadi *overfitting* atau *underfitting*. *Hyperparameter* yang digunakan biasanya berupa:

- *Epoch*

merupakan jumlah iterasi penuh yang dilakukan pada seluruh dataset pelatihan. Setiap *epoch* memungkinkan model untuk melihat seluruh data dan memperbarui bobotnya. Jumlah *epoch* yang optimal membantu model untuk belajar tanpa mengalami *overfitting* atau *underfitting*.

- *Batch Size*

merupakan jumlah sampel data yang digunakan dalam satu iterasi untuk memperbarui bobot model. Penggunaan *batch size* yang tepat dapat mempengaruhi kecepatan pelatihan dan kualitas pembelajaran. *Batch size* yang kecil dapat membuat model lebih fleksibel, sementara *batch size* yang lebih besar mempercepat pelatihan.

- *Optimizer*

Merupakan algoritma yang digunakan untuk memperbarui bobot model berdasarkan kesalahan yang dihitung selama pelatihan. Digunakan *optimizer* Adam dalam penelitian ini yang menggabungkan keuntungan dari momentum dan adaptif *learning rate* untuk membuat pelatihan lebih efisien.

- *Learning Rate*

Learning rate mengontrol seberapa besar pembaruan bobot yang dilakukan selama pelatihan. Nilai *learning rate* yang terlalu besar dapat menyebabkan model melompati solusi optimal, sementara nilai yang terlalu kecil dapat membuat pelatihan terlalu lambat dan mempengaruhi konvergensi model.

Proses pelatihan model LSTM dimulai dengan membagi dataset menjadi dua bagian: data latih dan data uji. Data latih digunakan untuk melatih model dalam mempelajari pola-pola dalam urutan kata, sedangkan data uji digunakan untuk mengukur kemampuan model dalam mengklasifikasi sentimen yang belum pernah dilihat sebelumnya. Selama pelatihan, input berupa urutan kata dari tweet diproses melalui lapisan LSTM untuk menangkap dependensi jangka panjang antar kata. Setiap lapisan LSTM akan memperbarui cell state dan menghasilkan output berdasarkan informasi yang relevan dari kata-kata sebelumnya. Setelah model dilatih pada data latih, model kemudian diuji menggunakan data uji untuk memverifikasi seberapa baik model mengklasifikasi sentimen (positif atau negatif).

Proses pelatihan dilakukan dalam beberapa *epoch*, dengan penggunaan *batch size* yang telah ditentukan untuk memastikan pembaruan bobot dilakukan secara optimal.

G. Evaluasi dan Analisis

Setelah proses pelatihan, model LSTM dievaluasi menggunakan data uji untuk mengukur kinerjanya dalam mengklasifikasikan sentimen publik terhadap kebijakan Efisiensi Anggaran 2025. Beberapa metrik yang digunakan dalam evaluasi ini adalah akurasi, presisi, *recall*, dan *F1-score*. Akurasi mengukur proporsi tweet yang diklasifikasikan dengan benar, sementara precision dan *recall* mengevaluasi kemampuan model dalam mengenali tweet dengan sentimen positif dan negatif. *F1-score* memberikan gambaran keseimbangan antara precision dan *recall*.

Selain itu, *confusion matrix* digunakan untuk menilai distribusi hasil klasifikasi dan mengidentifikasi kesalahan yang terjadi, seperti *false positive* dan *false negative*. Dengan menggunakan metrik tersebut, hasil evaluasi memberikan pemahaman mendalam mengenai keefektifan model LSTM dalam memprediksi sentimen dari tweet yang terkait dengan kebijakan tersebut. Berikut merupakan contoh dari *confusion matrix*.

True Labels	Actual Values	
	0 (Negatif)	1 (Positif)
0 (Negatif)	TN	FP
1 (Positif)	FN	TP

Gambar. 4 Confusion Matrix

Analisis juga dilakukan terhadap keseluruhan data untuk menilai respon masyarakat terhadap kebijakan Efisiensi Anggaran 2025. Analisis ini bertujuan untuk menggali pola-pola sentimen yang muncul dari setiap tweet yang ada, baik yang mendukung maupun yang menentang kebijakan tersebut. Hasil analisis ini diharapkan dapat memberikan wawasan yang berguna bagi pemerintahan dalam mengevaluasi kebijakan yang ada, serta menjadi dasar pertimbangan dalam merumuskan kebijakan lainnya di masa depan.

HASIL DAN PEMBAHASAN

Dilakukan beberapa eksperimen dengan berbagai rasio pembagian data untuk pelatihan dan pengujian yaitu rasio 70:30, 80:20, 90:10. Pengaturan *hyperparameter* juga dilakukan yaitu dengan *epoch* 10,20,30 serta *batch size* 32 dan 64 yang masing-masing akan di uji coba. Hal ini

dilakukan agar dapat mengetahui seberapa besar pengaruh terhadap akurasi dan kinerja model LSTM dalam mengklasifikasikan sentimen publik terkait kebijakan Efisiensi Anggaran 2025.

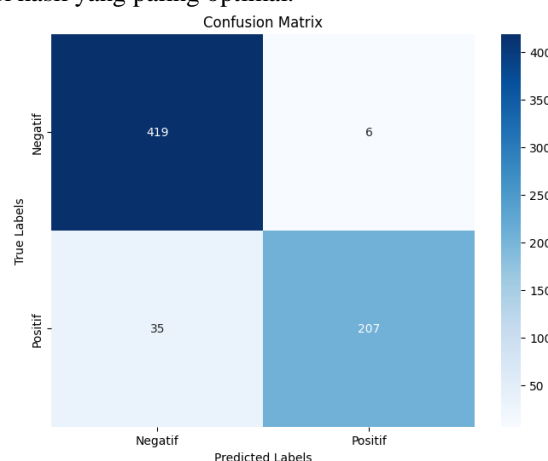
A. Uji Coba Rasio 70:30

Dimulai dari rasio 70:30 dengan menggunakan 1555 data untuk pelatihan serta 667 data untuk pengujian. Proses pelatihan melibatkan *hyperparameter epoch* serta *batch size* untuk mengoptimalkan pembelajaran model. Hasil pengujian serta evaluasi pada rasio ini adalah sebagai berikut:

Tabel. 4 Hasil Uji Coba Rasio 70:30

Epoch	Batch Size	Akurasi	Presisi	Recall	F1-Score
10	32	95.20%	97.41%	87.44%	92.16%
20	32	94.90%	98.40%	85.58%	91.54%
30	32	93.70%	92.20%	87.91%	90.00%
10	64	94.60%	94.97%	87.91%	91.30%
20	64	91.75%	91.75%	87.91%	89.79%
30	64	90.78%	90.78%	86.98%	88.84%

Melihat hasil dari **Tabel 4** hasil terbaik dari uji coba pada rasio ini terdapat pada hasil yang memperoleh nilai akurasi sebesar 95,20% dengan presisi 97,41%, *recall* 87,44%, dan *F1-score* 92,16% pada epoch 10 dan batch size 32. Meskipun terdapat peningkatan dengan batch size 64, model mengalami penurunan performa pada beberapa metrik ketika jumlah *epoch* ditingkatkan. Hal ini menunjukkan bahwa pada rasio 70:30, model dapat mencapai kinerja yang baik pada data pelatihan terbatas, tetapi semakin banyak *epoch* dapat menyebabkan sedikit penurunan performa. Berikut evaluasi *confusion matrix* dari hasil yang paling optimal.



Gambar. 5 Confusion Matrix Rasio 70:30

Gambar 5 di atas menunjukkan *confusion matrix* yang digunakan untuk mengevaluasi hasil klasifikasi sentimen oleh model LSTM. Pada matrix ini, terdapat 419 tweet yang benar-benar diklasifikasikan sebagai negatif (*True Negatif*), dan 207 tweet yang benar-benar diklasifikasikan sebagai positif (*True Positif*). Namun, terdapat beberapa kesalahan klasifikasi, yaitu 6 tweet yang sebenarnya

negatif, tetapi salah diklasifikasikan sebagai positif (*False Positif*), serta 35 tweet yang sebenarnya positif, tetapi salah diklasifikasikan sebagai negatif (*False Negatif*). Secara keseluruhan, *confusion matrix* ini menunjukkan bahwa model LSTM memiliki kinerja yang cukup baik dengan jumlah kesalahan klasifikasi yang relatif rendah, meskipun masih ada ruang untuk perbaikan terutama dalam mengurangi *false negatives*.

B. Uji Coba Rasio 80:20

Pada uji coba dengan rasio 80:20, yang berarti 80% data digunakan untuk pelatihan dan 20% untuk pengujian, model LSTM menunjukkan peningkatan kinerja yang signifikan dibandingkan dengan rasio 70:30. Dengan menggunakan 1778 data untuk pelatihan dan 444 data untuk pengujian. Hasil dari pengujian rasio tersebut terdapat dalam tabel berikut:

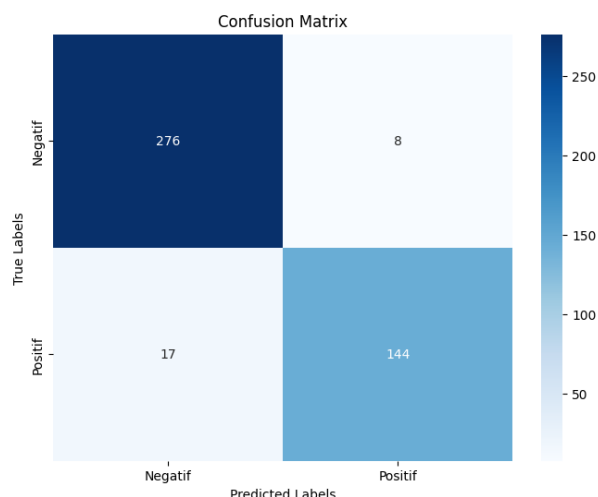
Tabel. 5 Hasil Uji Coba Rasio 80:20

Epoch	Batch Size	Akurasi	Presisi	Recall	F1-Score
10	32	94.83%	97.41%	84.96%	90.76%
20	32	93.93%	93.44%	85.71%	89.41%
30	32	93.48%	91.27%	86.47%	88.80%
10	64	94.16%	97.35%	82.71%	89.43%
20	64	93.48%	90.62%	87.22%	88.89%
30	64	93.71%	92.00%	86.47%	89.15%

Dari hasil **Tabel 5**, terlihat bahwa model dengan *epoch* 10 dan *batch size* 32 menghasilkan performa terbaik dengan akurasi 94,38% dan *F1-score* 97,41%. Rasio 80:20 memberikan model lebih banyak data untuk pelatihan, yang memungkinkan model LSTM untuk lebih baik dalam mengenali pola sentimen positif meskipun ada tantangan dalam klasifikasi sentimen negatif.

Namun, meskipun hasil presisi meningkat, *recall* untuk sentimen negatif masih menunjukkan beberapa kesalahan klasifikasi. Ini menunjukkan bahwa meskipun model dapat mengenali pola positif dengan baik, masih terdapat kesulitan dalam menangani sentimen negatif dengan akurat. Dengan *batch size* 64, meskipun presisi sedikit meningkat, *recall* mengalami penurunan yang menunjukkan bahwa terlalu banyak data per batch dapat mengurangi efektivitas model dalam mengidentifikasi sentimen negatif.

Selanjutnya, **Gambar 6** menunjukkan *confusion matrix* yang *epoch* 10 dan *batch size* 32 pada rasio 80:20 yang digunakan untuk mengevaluasi hasil klasifikasi sentimen model LSTM. Pada matrix ini, terdapat 276 tweet negatif yang diklasifikasikan dengan benar (*True Negatif*) dan 144 tweet positif yang benar (*True Positif*). Namun, terdapat kesalahan klasifikasi dengan 8 tweet negatif yang salah diklasifikasikan sebagai positif (*False Positif*) dan 17 tweet positif yang salah diklasifikasikan sebagai negatif (*False Negatif*).



Gambar. 6 Confusion Matrix Rasio 80:20

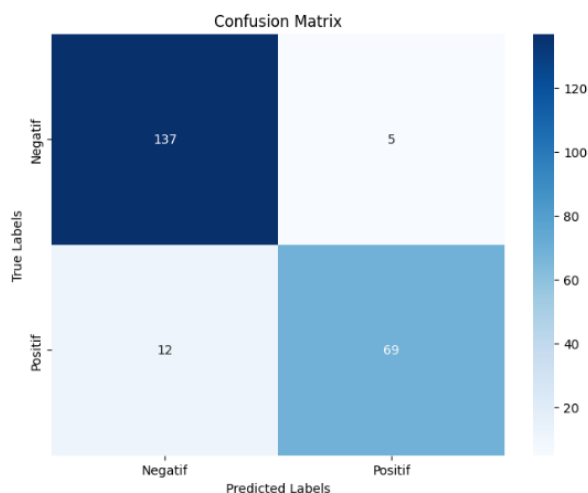
C. Uji Coba Rasio 90:10

Pada uji coba dengan rasio 90:10, di mana 90% data digunakan untuk pelatihan dan 10% untuk pengujian, model LSTM menunjukkan hasil yang baik, meskipun sedikit menurun dibandingkan dengan rasio sebelumnya. Dengan menggunakan 1999 data untuk pelatihan dan 223 data untuk pengujian. Tabel berikut akan menampilkan table untuk hasil pengujian pada rasio ini.

Tabel. 6 Hasil Uji Coba Rasio 90:10

Epoch	Batch Size	Akurasi	Presisi	Recall	F1-Score
10	32	94.62%	93.85%	88.41%	91.04%
20	32	93.17%	92.42%	88.41%	90.37%
30	32	92.62%	95.24%	86.96%	90.20%
10	64	94.17%	95.16%	85.51%	90.51%
20	64	92.93%	94.37%	82.72%	88.16%
30	64	91.48%	90.79%	85.19%	87.90%

Dari hasil **Tabel 6** tersebut, dapat model model berhasil mencapai akurasi 94,62%. Metrik presisi dan *recall* pada rasio ini masing-masing mencapai 93,85% dan 88,41%, dengan *F1-score* sebesar 90,04% pada epoch 10 dan batch size 32. Walaupun hasil akurasi cukup tinggi, terdapat sedikit penurunan pada metrik *recall* dan precision dibandingkan dengan rasio 80:20, yang menunjukkan bahwa model mulai mengalami kesulitan dalam mengidentifikasi sentimen negatif ketika lebih sedikit data digunakan untuk pengujian. Evaluasi lebih lanjut dengan *confusion matrix* akan memberikan gambaran lebih mendalam mengenai kesalahan klasifikasi yang terjadi.



Gambar. 7 Confusion Matrix Rasio 90:10

Gambar 7 menunjukkan confusion matrix yang digunakan untuk mengevaluasi hasil klasifikasi sentimen oleh model LSTM pada rasio 90:10. Dalam matrix ini, terdapat 137 tweet negatif yang diklasifikasikan dengan benar (*True Negatif*) dan 69 tweet positif yang diklasifikasikan dengan benar (*True Positif*). Namun, terdapat juga beberapa kesalahan klasifikasi, yaitu 5 tweet negatif yang salah diklasifikasikan sebagai positif (*False Positif*) dan 12 tweet positif yang salah diklasifikasikan sebagai negatif (*False Negatif*). Secara keseluruhan, model LSTM memiliki kinerja yang baik meskipun ada beberapa kesalahan dalam klasifikasi sentimen, terutama pada tweet dengan sentimen positif yang lebih sedikit diklasifikasikan dengan benar. Perbaikan dalam menangani false negatives dapat meningkatkan kinerja model lebih lanjut.

D. Analisis Konten

Penelitian ini mengkaji Kebijakan Efisiensi Anggaran 2025 dengan menganalisis 2222 *tweet* yang diambil dari platform X. Dari total *tweet* tersebut, 806 *tweet* berlabel positif dan 1416 *tweet* berlabel negatif. Analisis dilakukan untuk mengidentifikasi kriteria yang mendasari setiap label, serta masalah yang diungkapkan masyarakat terkait kebijakan ini.

• Tweet Kategori Negatif

Label negatif mencerminkan pandangan masyarakat yang mengkhawatirkan dampak pengurangan anggaran. Beberapa isu yang diangkat dalam tweet negatif adalah:

1. Kekhawatiran Pemangkasan Anggaran pada Layanan Publik
Banyak yang khawatir dengan pengurangan dana untuk sektor vital seperti pendidikan dan kesehatan. Seperti *tweet* yang menyebutkan, "Efisiensi anggaran bisa mengurangi dana untuk sektor penting seperti pendidikan."
2. Penggunaan Anggaran untuk Kepentingan Pejabat
Beberapa *tweet* mengkritik penggunaan anggaran yang tidak transparan dan digunakan untuk

kepentingan pribadi pejabat, "Efisiensi anggaran malah jadi alasan buat pejabat korup."

3. Pengurangan Pendapatan PNS

Kebijakan ini juga mempengaruhi kesejahteraan Pegawai Negeri Sipil (PNS), seperti yang tercatat dalam *tweet*, "Gaji PNS bisa dipotong hanya demi efisiensi?"

4. Kesulitan Masyarakat Akibat Penghematan

Banyak yang merasa kebijakan ini memberi dampak negatif langsung pada mereka, terutama dalam hal pengurangan program sosial, seperti yang disampaikan, "Efisiensi anggaran hanya bikin masyarakat kesulitan."

Secara keseluruhan, tweet negatif mengungkapkan kekhawatiran masyarakat terhadap transparansi pengelolaan anggaran dan dampak pengurangan dana pada sektor-sektor vital serta kesejahteraan publik.

• Tweet Kategori Positif

Label positif mencerminkan pandangan masyarakat yang mendukung kebijakan Efisiensi Anggaran 2025 sebagai langkah yang dapat membawa dampak positif. Beberapa masalah yang diidentifikasi dari *tweet* positif antara lain:

1. Penyediaan Informasi yang Jelas

Masyarakat menginginkan informasi yang jelas untuk mengurangi kekhawatiran terkait kebijakan ini. Seperti yang terlihat dalam *tweet*, "Anggaran efisiensi tidak sama dengan anggaran CPNS, perlu penjelasan lebih lanjut."

2. Penyesuaian Anggaran yang Responsif

Perubahan kondisi membutuhkan penyesuaian anggaran agar tetap sesuai dengan rencana, seperti yang diungkapkan dalam *tweet*, "Kondisi selalu berubah, justru bagus kalau anggaran tetap berjalan sesuai rencana."

3. Efisiensi Sebagai Langkah Penyusunan Ulang Prioritas

Beberapa pihak melihat kebijakan ini sebagai cara untuk memastikan anggaran digunakan secara efektif, seperti *tweet* yang menyatakan, "Efisiensi anggaran dapat memperbaiki alokasi dana yang sebelumnya tidak tepat."

4. Menjaga Kelangsungan Program Pemerintah

Efisiensi anggaran dianggap penting untuk menjaga kelangsungan program pemerintah, meskipun terjadi penghematan dana, seperti yang disebutkan, "Meskipun terjadi pengurangan anggaran, program penting tetap berjalan."

Kebijakan ini dipandang sebagai langkah positif dalam pengelolaan keuangan negara yang lebih tepat sasaran dan berfokus pada prioritas yang mendesak.

KESIMPULAN DAN SARAN

Penelitian ini mengkaji pengaruh metode LSTM dalam mengklasifikasikan sentimen publik terkait Kebijakan Efisiensi Anggaran 2025 menggunakan dataset 2222 *tweet* yang telah dibersihkan dan dilabeli sebagai sentimen

positif dan negatif. Hasil penelitian menunjukkan bahwa LSTM dan TF-IDF efektif dalam memproses data ini dan memberikan akurasi yang baik pada tiga rasio pembagian data: 70:30, 80:20, dan 90:10. Akurasi tertinggi tercatat pada rasio 80:20, yaitu 94,38%, dengan *epoch* 10 dan *batch size* 32, yang menunjukkan bahwa model bekerja optimal dengan pembagian data yang lebih seimbang. Ini mengindikasikan bahwa pengaturan *hyperparameter* yang tepat dan pembagian data yang sesuai sangat penting dalam mencapai kinerja terbaik pada model LSTM.

Untuk penelitian selanjutnya, disarankan untuk menyeimbangkan jumlah data kategori sentimen agar mengurangi bias dalam model, dengan menggunakan metode seperti *oversampling* atau SMOTE. Selain itu, penggunaan metode pembobotan kata lain seperti Word2Vec atau GloVe dapat lebih memahami konteks kalimat dibandingkan TF-IDF. Peneliti juga dapat mencoba metode yang lebih canggih, seperti GRU atau Transformer, yang lebih efisien dalam menangani data sekuensial panjang.

Berdasarkan temuan analisis sentimen, sebagian besar tanggapan negatif dari masyarakat dapat diatasi dengan memperjelas alokasi anggaran untuk sektor penting, meningkatkan pengawasan penggunaan anggaran, serta memastikan kesejahteraan PNS tanpa memberatkan masyarakat. Penyuluhan untuk meningkatkan pemahaman publik dan transparansi dalam pengalokasian dana juga dapat membantu memperbaiki respons masyarakat terhadap kebijakan ini.

REFERENSI

- [1] S. Lestari and Z. A. Suarja, "Blue Ekonomi: Dampak Perencanaan Anggaran Terhadap Pembangunan Ekonomi Berkelanjutan," *Journal of Economics Science*, vol. 9, no. 1, doi: 10.33143/jecs.v9i1.3003.
- [2] H. Walizi, "Dampak Kebijakan Efisiensi Anggaran Terhadap Pencapaian Sasaran Pembangunan Daerah Di BAPPEDA MUARA ENIM," *Integrative Perspectives of Social and Science Journal*, vol. 2, no. 2, pp. 1707–1718.
- [3] Badan Pemeriksa Keuangan, "Efisiensi Belanja dalam Pelaksanaan Anggaran Pendapatan dan Belanja Negara dan Anggaran Pendapatan dan Belanja Daerah Tahun Anggaran 2025," peraturan.bpk.go.id. [Online]. Available: <https://peraturan.bpk.go.id/Details/313401/inpres-no-1-tahun-2025>
- [4] R. A. Huda, R. A. Puspita, S. N. Hasanah, and A. Malik, "Peran Kebijakan Fiskal Dalam Meningkatkan Kesejahteraan Sosial Di Indonesia Sebagai Negara Berkembang," *KALIANDA HALOK GAGAS*, vol. 7, no. 2, 2024.
- [5] A. F. Ilma and A. I. Prabandari, "Efisiensi vs. Kualitas: Evaluasi Kebijakan Anggaran Pendidikan dalam Pemerintahan Prabowo," *Lingkar Ekonomika*, vol. 4, no. 1, doi: 10.32424/jle.v4i1.15823.
- [6] A. Rahman, "Analisis Drone Empirit," Analisis Percakapan Publik: Efisiensi Anggaran. [Online]. Available: <https://pers.droneempit.id/analisis-percakapan-publik-efisiensi-anggaran/>
- [7] Restu Aprianto Atmiko, Melaya Priangga, Nabil Panca Budhi Hizbullah, Muhammad Anis Nur Fauzi, and Dzulfan Yumna Azis, "Pemanfaatan Data X Untuk Analisis Sentimen Publik Pada Kebijakan PPN 12%," *JITE*, vol. 2, no. 1, pp. 274–283, Apr. 2025, doi: 10.63547/jite.v2i1.55.
- [8] S. Sudriyanto, M. Faid, K. Malik, and A. Supriadi, "Evaluasi Model Jaringan Saraf Tiruan Berbasis LSTM dalam Memprediksi Fluktuasi Harga Bitcoin," *JARS*, vol. 2, no. 2, pp. 15–22, June 2024, doi: 10.24929/jars.v2i2.3398.
- [9] Nailah Azzahra, Merry Dwi Handayani, and Awwaliyah Aliyah, "Evaluasi Kinerja AI berbasis Recurrent Neural Network (RNN) dalam Mengidentifikasi Ancaman Phising pada URL Website," *Bridge*, vol. 3, no. 3, pp. 15–37, June 2025, doi: 10.62951/bridge.v3i3.485.
- [10] W. P. Anggraini and M. S. Utami, "Klasifikasi Sentimen Masyarakat Terhadap Kebijakan Kartu Pekerja Di Indonesia," *Faktor Exacta*, vol. 13, no. 4, pp. 255–255, 2021, doi: 10.30998/faktorexacta.v13i4.7964.
- [11] F. Mahmuji, C. Muchammad, C. Cahyo, and U. Nisa, "Analisis Sentimen Pada Twitter Terhadap Isu Penundaan Pemilu 2024 Dengan Membandingkan Metode Long Short-Term Memory Dan Naive Bayes Classifier," *Journal of Mathematics & Information Technology*, vol. 1, no. 2, 2023, doi: 10.35718/equiva.v1i2.
- [12] A. F. Aufar, Mochamad Alfian Rosid, A. Eviyanti, and I. R. I. Astutik, "Optimizing Text Preprocessing for Accurate Sentiment Analysis on E-Wallet Reviews," *JICTE*, vol. 7, no. 2, pp. 42–50, Oct. 2023, doi: 10.21070/jicte.v7i2.1650.
- [13] A. A. Firdaus, "Sentiment Budget Efficiency in Indonesia." 2025. Accessed: June 16, 2025. [Online]. Available: <https://www.kaggle.com/datasets/jocelyndumlao/budget-efficiency-in-indonesia>
- [14] N. I. Widiastuti, E. Rainarli, and K. E. Dewi, "Peringkasan dan Support Vector Machine pada Klasifikasi Dokumen," *INFOTEL*, vol. 9, no. 4, p. 416, Nov. 2017, doi: 10.20895/infotel.v9i4.312.
- [15] M. K. Wisyaldin, G. M. Luciana, and H. Pariaman, "Pendekatan Long Short-Term Memory untuk Memprediksi Kondisi Motor 10 kV pada PLTU Batubara," *KILAT*, vol. 9, no. 2, 2020, doi: 10.33322/kilat.v9i2.997.