Komputika: Jurnal Sistem Komputer
Volume 14, Nomor 2, Oktober 2025, hlm. 193 - 202
DOI: 10.34010/komputika.v14i2.16976

p-ISSN: 2252-9039
e-ISSN: 2655-3198

Analisis Sentimen Pengguna Twitter Mengenai Kotak Kosong di Pilkada Indonesia Tahun 2024 Menggunakan Algoritma LSTM

Achmad Chikham Nouriel Rosyadi[1*], Muhammad Athoillah[2], Fenny Fitriani[3]
[1,2,3)] Program Studi Statistika, Fakultas Teknik dan Sains, Universitas PGRI Adi Buana Surabaya
Jl. Dukuh Menanggal XII, Dukuh Menanggal, Kec. Gayungan, Surabaya, Jawa Timur 60234

email: chikam015@gmail.com

(Naskah masuk: *14 Juli 2025*; direvisi: *16 September 2025*; diterima untuk diterbitkan: *02 Oktober 2025*)

ABSTRAK *– Fenomena kotak kosong dalam Pilkada 2024 menjadi perbincangan ramai di Twitter, memunculkan pandangan pro dan kontra di masyarakat mengenai keberadaan calon tunggal yang dinilai berpotensi memengaruhi stabilitas politik di daerah. Diskusi publik di media sosial umumnya menggunakan bahasa informal, dialek bahasa daerah, serta tidak mengikuti kaidah baku, sehingga menyulitkan analisis manual dan berisiko menimbulkan bias. Maka dari itu, penelitian ini mengimplementasikan pendekatan deep learning dengan memanfaatkan model Long Short-Term Memory (LSTM) dan word embedding GloVe. Proses pelabelan dilakukan secara otomatis menggunakan Indonesia Sentiment Lexicon (INSET) untuk mengklasifikasikan sentimen masyarakat terhadap kebijakan kotak kosong berdasarkan data Twitter. Data penelitian terdiri dari 2.168 tweet yang diperoleh melalui teknik crawling, kemudian dievaluasi menggunakan metode 10-fold cross-validation. Analisis sentimen menghasilkan distribusi opini publik, yaitu 35,9% negatif, 32,8% positif, dan 31,3% netral. Hasil pengujian menunjukkan akurasi tertinggi sebesar 94,93% pada fold ke-6, dengan rata-rata akurasi keseluruhan mencapai 90,08%. Penelitian ini berkontribusi dalam pengembangan sistem analisis berbasis deep learning untuk pemantauan opini publik, serta menunjukkan potensi data media sosial sebagai sumber informasi strategis dalam memahami persepsi masyarakat terhadap fenomena politik.*

Kata Kunci *– Deep Learning; Demokrasi; GloVe; Opini publik; Text Mining.*

SENTIMENT ANALYSIS OF TWITTER USERS REGARDING THE EMPTY BOX IN INDONESIA'S 2024 REGIONAL HEAD ELECTION USING THE LSTM ALGORITHM

ABSTRACT *– The empty ballot box phenomenon in the 2024 Regional Head Elections (Pilkada) has become a widely discussed topic on Twitter, sparking both pro and contra views among the public regarding the existence of a single candidate, which is considered to have the potential to affect political stability at the regional level. Public discussions on social media generally use informal language, regional dialects, and do not follow formal linguistic rules, making manual analysis difficult and potentially biased. Therefore, this study implements a deep learning approach by utilizing the Long Short-Term Memory (LSTM) model and GloVe word embedding. The labeling process was conducted automatically using the Indonesia Sentiment Lexicon (INSET) to classify public sentiment toward the empty ballot box policy based on Twitter data. The research dataset consists of 2,168 tweets collected through a crawling technique, then evaluated using the 10-fold cross-validation method. Sentiment analysis resulted in the distribution of public opinion: 35.9% negative, 32.8% positive, and 31.3% neutral. The testing results show the highest accuracy of 94.93% on the 6th fold, with an overall average accuracy of 90.08%. This study contributes to the development of deep learning-based analysis systems for monitoring public opinion and demonstrates the potential of social media data as a strategic source of information in understanding public perceptions of political phenomena.*

Keywords *– Deep Learning; Democracy; GloVe; Public Opinion; Text Mining.*

1. PENDAHULUAN

Pemilihan Kepala Daerah (Pilkada) merupakan salah satu bentuk demokrasi langsung di Indonesia yang memberikan ruang partisipasi politik bagi masyarakat dalam menentukan pemimpin di tingkat daerah [1]. Pilkada diselenggarakan secara serentak setiap lima tahun, dan merupakan pilar penting

dalam sistem desentralisasi pemerintahan. Dalam praktiknya, pelaksanaan Pilkada tidak selalu diwarnai dengan kehadiran lebih dari satu pasangan calon. Pada kondisi tertentu, hanya terdapat satu pasangan calon yang maju, sehingga berdasarkan Undang-Undang Nomor 10 Tahun 2016 Pasal 54C dan 54D, masyarakat tetap diberikan dua pilihan, yakni mencoblos pasangan calon tersebut atau mencoblos kotak kosong. Jika kotak kosong memperoleh suara lebih dari 50%, maka Pilkada harus diulang.

Kebijakan kotak kosong dalam Pilkada menimbulkan pro dan kontra di tengah masyarakat dan pengamat politik. Di satu sisi, dianggap sebagai representasi dari ketidakpuasan terhadap calon tunggal yang diusung partai politik. Di sisi lain, kebijakan ini juga dianggap dapat mengganggu stabilitas politik dan menciptakan ketidakpastian pemerintahan di tingkat daerah, terutama ketika terjadi kekosongan kepemimpinan akibat hasil Pilkada yang dimenangkan oleh kotak kosong [2]. Pro dan kontra kotak kosong pilkada ini juga menjadi topik utama pada sosial media twitter, dimana twitter dengan pengguna sosial media berbasis teks terbanyak [3]. Namun karakter teks yang digunakan dalam media sosial twitter cenderung informal, singkat, mengandung bahasa gaul, hingga dialek lokal, sehingga menyulitkan proses analisis manual, terlebih dalam jumlah data yang sangat besar, sehingga membutuhkan pendekatan berbasis *text mining* dan *Natural Language Processing* (NLP). Perkembangan sistem otomatis untuk klasifikasi sentimen di media sosial telah menjadi semakin canggih, seiring dengan kemajuan dalam bidang *machine learning* dan NLP [4]. *Deep learning* merupakan perkembangan dari *machine learning*, memungkinkan untuk melakukan klasifikasi sentimen dengan otomatis dan hasil yang lebih baik pada data teks [5].

Salah satu metode *deep learning* yang bisa digunakan dalam pemrosesan data teks adalah *Long Short-Term Memory* (LSTM). LSTM memiliki kemampuan dalam melakukan klasifikasi sentimen yakni proses berjalannya algoritma ini berurutan dan mempertimbangkan input sebelumnya, sehingga memiliki kemampuan memproses informasi sekuensial dan pola kalimat pada data teks [6]. Sebelum diterapkan ke dalam model, teks biasanya direpresentasikan dalam bentuk numerik menggunakan teknik *word embedding*, seperti *GloVe* (*Global Vectors for Word Representation*), yang menangkap makna semantik dan hubungan antar kata berdasarkan statistik kemunculan bersamaan dalam korpus besar [7].

Beberapa penelitian sebelumnya menunjukkan keberhasilan LSTM dalam klasifikasi teks dan analisis sentimen. Penelitian klasifikasi teks media sosial menggunakan *word embeddings* dan LSTM, dimana klasifikasi kecenderungan politik menghasilkan akurasi mencapai 87,57% [8]. Penelitian yang lain, mengembangkan model LSTM dengan *lexicon* dan *attention* untuk analisis sentimen teks umum. Model ini secara signifikan meningkatkan akurasi klasifikasi dibandingkan LSTM standar dengan *word embedding* [9]. Selain itu, pada penelitian perbandingan penyematan kata untuk analisis sentimen bahasa indonesia, model LSTM menunjukkan performa lebih baik dibanding CNN. *Word2Vec* dan *FastText* memiliki keunggulan masing-masing dalam menangkap konteks bahasa Indonesia. Secara umum, LSTM menunjukkan performa yang unggul dalam menangani bahasa alami, termasuk yang bersifat morfologis kompleks.

Berdasarkan penelitian terdahulu belum ada studi yang secara khusus menerapkan model LSTM dengan kombinasi *GloVe* pada teks berbahasa Indonesia, pada konteks kotak kosong Pilkada di indonesia dengan pelabelan menggunakan kamus Indonesia *Sentiment Lexicon* (INSET).

Dengan latar belakang tersebut, penelitian ini berfokus pada analisis sentimen media sosial twitter berbahasa Indonesia, dengan memanfaatkan kombinasi *word embedding GloVe* dan LSTM. Temuan dari studi ini diharapkan dapat dimanfaatkan oleh penyelenggara pemilu dan pembuat kebijakan sebagai landasan dalam menyusun strategi komunikasi publik, serta memperkuat kajian akademik terkait *deep learning*.

2. METODE DAN BAHAN

Rancangan Penelitian

Tahap awal pada penelitian ini dimulai dengan pengumpulan data. Didapatkan sebanyak 2.168 tweet yang membahas mengenai kotak kosong pada pilkada 2024. Pengambilan data Twitter dilakukan dengan bantuan google colab menggunakan *library tweet-harvest*, dapat dilihat pada Gambar 1 berikut.

```
# Crawl Data

filename = 'kotak_kosong.csv'
search_keyword = 'kosong kotak lang:id until:2024-12-31 since:2024-08-01'
limit = 5000

!npx -y tweet-harvest@2.6.1 -o "{filename}" -s "{search_keyword}" --tab "LATEST"
-l {limit} --token {twitter_auth_token}
```

Gambar 1. *Syntax Crawling* Data Twitter

Setelah proses pengumpulan data selesai dilakukan, tahap berikutnya adalah data *preprocessing*, yang mencakup proses *case folding*, yakni mengubah huruf teks menjadi huruf kecil seluruhnya, dengan menggunakan *syntax* yang ditunjukkan pada Gambar 2 berikut.

```
# ------ Case Folding --------
# Using Series.str.lower() function on Pandas
tweet['Text Case Folding'] = tweet['Text'].str.lower()


print('Case Folding Result : \n')
print(tweet['Text Case Folding'].head(5))
print('\n\n\n')
```

Gambar 2. *Syntax Crawling* Data Twitter

Selanjutnya melakukan penghapusan url dan tanda baca serta *tokenizing*, yakni mengubah tweet menjadi potongan kata pada dataset twitter. Menggunakan *syntax* yang ditunjukkan pada Gambar 3 berikut.

```
# ------ Tokenizing ---------

def remove_tweet_special(text):
    # remove tab, new line, ans back slice
    text = text.replace('\\t'," ").replace('\\n'," ").replace('\\u'," ").replace('\\',"")
    # remove non ASCII (emoticon, chinese word, .etc)
    text = text.encode('ascii', 'replace').decode('ascii')
    # remove mention, link, hashtag
    text = ' '.join(re.sub("([@#][A-Za-z0-9]+)|(\w+:\/\/\S+)"," ", text).split())
    # remove incomplete URL
    return text.replace("http://", " ").replace("https://", " ")

tweet['Text Case Folding'] = tweet['Text Case Folding'].apply(remove_tweet_special)

#remove number
def remove_number(text):
    return  re.sub(r"\d+", "", text)

tweet['Text Case Folding'] = tweet['Text Case Folding'].apply(remove_number)

#remove punctuation
def remove_punctuation(text):
    return text.translate(str.maketrans("","",string.punctuation))

tweet['Text Case Folding'] = tweet['Text Case Folding'].apply(remove_punctuation)
```

Gambar 3. *Syntax* Penghapusan Tanda Baca dan *Tokenizing* Data Twitter

Tahap selanjutnya adalah normalisasi *slang word*, yakni merubah kata atau istilah gaul (*slang*) menjadi kata baku, dengan menggunakan *syntax* yang ditunjukkan pada Gambar 4.

```
# ========== NORMALIZATION ==========
import json

# Bangun dict dari bentuk slang ke bentuk baku
def load_slang_dict():
    with open('slang_word_dict.txt', 'r') as file:
        slang_dict_raw = json.load(file)  # lebih aman daripada json.loads(file.read())
    slang_to_baku = {}
    for baku, slang_list in slang_dict_raw.items():
        for slang in slang_list:
            slang_to_baku[slang.lower()] = baku
    return slang_to_baku

slang_to_baku = load_slang_dict()

def normalize_slang_words(tokenized_sentence):
    return [slang_to_baku.get(token.lower(), token) for token in tokenized_sentence]
tweet['Text Normalization'] = tweet['Text Tokenizing'].apply(normalize_slang_words)
print("Normalization completed")
```

Gambar 4. *Syntax* Normalisasi *Slang Word* Data Twitter

Tahap selanjutnya adalah penghapusan *stopwords*, yakni menghapus kata dalam tweet tidak memiliki makna emosi. Menggunakan *syntax* yang ditunjukkan pada Gambar 5.

```
# ---------------------- add stopword from txt file ------------------------------
# read txt stopword using pandas
txt_stopword = pd.read_csv("stopwordbahasa.txt", names= ["stopwords"], header = None)

# convert stopword string to list & append additional stopword
list_stopwords.extend(txt_stopword["stopwords"][0].split(' '))

# ----------------------------------------------------------------------------------

# convert list to dictionary
list_stopwords = set(list_stopwords)


#remove stopword pada list token
def stopwords_removal(words):
    return [word for word in words if word not in list_stopwords]

tweet['Text Filtering'] = tweet['Text Stemming'].apply(stopwords_removal)


print(tweet['Text Filtering'].head())
```

Gambar 5. *Syntax* Penghapusan *Stopwords Data Twitter*

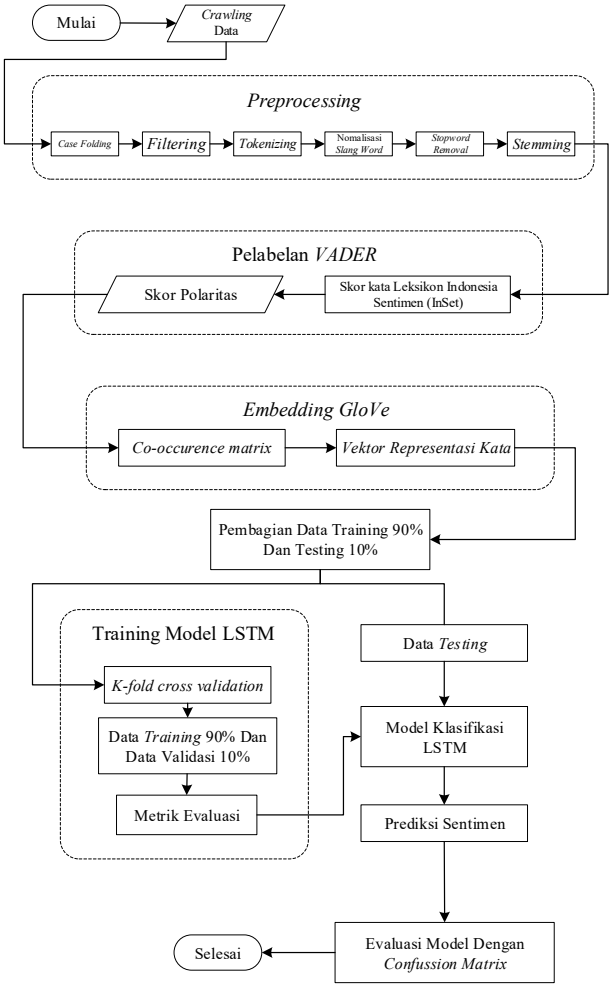
Tahap selanjutnya adalah *stemming*, yakni mengembalikan kata dalam twitter ke bentuk dasarnya, dengan menggunakan *syntax* yang ditunjukkan pada Gambar 6.

```
# create stemmer
factory = StemmerFactory()
stemmer = factory.create_stemmer()

# stemmed
def stemmed_wrapper(term):
    return stemmer.stem(term)

term_dict = {}

for document in tweet['Text Normalization']:
    for term in document:
        if term not in term_dict:
            term_dict[term] = ' '

print(len(term_dict))
print("------------------------")
```

Gambar 6. *Syntax Stemming* Data Twitter

Setelah data dibersihkan dan diproses, dilakukan visualisasi menggunakan *word cloud* untuk melihat kata-kata yang paling sering muncul, kemudian dilanjutkan dengan analisis sentimen menggunakan algoritma LSTM. Algoritma ini merupakan salah satu metode *deep learning[10]*. Setelah itu, proses dilanjutkan melakukan pelabelan sentimen positif, negatif dan netral menggunakan metode Indonesia *Sentiment Lexicon* (Inset), data kemudian dibagi menjadi dua bagian, yaitu data *training* dan data *testing*. Data *training* digunakan untuk melatih model dalam mengenali pola data dan melakukan klasifikasi kelas sentimen. Model yang dihasilkan dari tahap pelatihan ini diharapkan mampu merepresentasikan klasifikasi yang mendekati kondisi sebenarnya. Sementara itu, data *testing* berfungsi untuk mengevaluasi dan memvalidasi performa keandalan model yang telah dilatih. Proporsi pembagian data yang digunakan oleh peneliti adalah 90% untuk data *training* dan 10% untuk data *testing*. Model yang dibangun dengan menggunakan LSTM kemudian divalidasi menggunakan data *testing* untuk tingkat akurasi klasifikasi. Untuk memvalidasi model pada penelitian ini menggunakan metode *K-fold cross-*

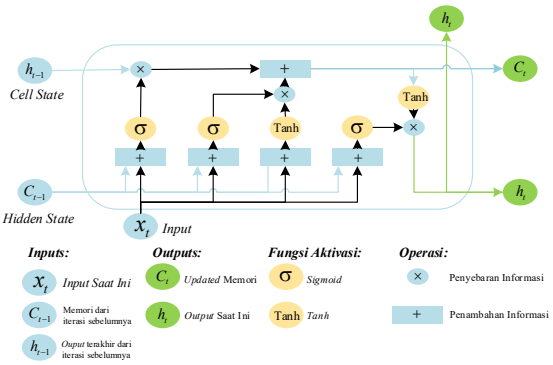
validation, yang akan menghasilkan nilai akurasi, presisi, *recall*, dan *F1-score* pada tiap kelas sentimen. Alur proses penelitian ini dijelaskan secara detail pada Gambar 7 berikut.



Gambar 7. Diagram Alir Klasifikasi LSTM

Long Short-Term Memory (LSTM)

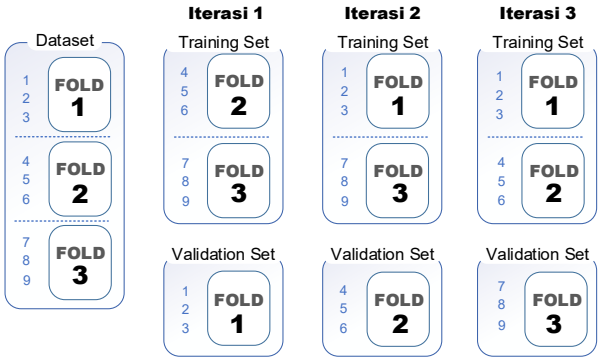
Long Short-Term Memory (LSTM) dikembangkan pada tahun 1997 untuk mengatasi kelemahan RNN dalam memproses data berurutan, khususnya dalam mempertahankan informasi jangka panjang [10]. RNN sering mengalami *vanishing gradient*, yang menghambat pelatihan pada data urutan panjang [11]. LSTM mengatasi masalah ini dengan menambahkan mekanisme gerbang seperti *input gate* dan *forget gate*, yang mengatur aliran informasi dan memungkinkan pemeliharaan konteks jangka panjang secara lebih efektif [12]. Pada gambar 8 terdapat arsitektur model LSTM.



Gambar 8. Arsitektur LSTM

K-fold Cross-Validation

Metode ini digunakan ketika jumlah data terbatas, dengan cara membagi dataset menjadi *k* bagian [13]. Proses k-*fold cross-validation* yakni membagi data menjadi k bagian yang memiliki ukuran sama atau hampir sama. Proses validasi kemudian dilakukan sebanyak k kali, di mana pada setiap iterasi, satu bagian digunakan sebagai data validasi, sedangkan k-1 lainnya digunakan sebagai data *training* untuk melatih model [14]. Kinerja model kemudian dihitung berdasarkan rata-rata hasil evaluasi dari seluruh iterasi tersebut. Gambar 9 berikut adalah ilustrasi dari *k-fold cross-validation* ketika k=3.



Gambar 9. Ilustrasi *3-Fold Cross-Validation*

3. HASIL DAN PEMBAHASAN

Hasil *Crawling* Data

Tahap awal dalam penelitian ini dimulai dengan pengambilan data dari Twitter menggunakan library *tweet-harvest*. Pengambilan data dilakukan berdasarkan kata kunci yang dirancang untuk menjaring percakapan relevan terkait kebijakan kotak kosong. Proses *crawling* dilakukan selama periode september hingga desember 2024, yang mencakup fase kampanye, hari pemilihan, dan pengumuman hasil Pilkada. Dari proses tersebut, berhasil dikumpulkan sebanyak 2.168 tweet.

Hasil *Preprocessing*

Tahap *preprocessing* dilakukan untuk membersihkan data dari karakter yang tidak diperlukan dan mengubah struktur teks menjadi format yang sesuai untuk pemodelan. Langkah ini juga bertujuan mengurangi *noise* agar model dapat lebih fokus mengenali pola. Selain itu, dilakukan penghapusan tweet duplikat guna mencegah bias dalam pelatihan, dengan total 2.168 tweet hasil *crawling*. Tabel 1 menunjukan hasil dari *preprocessing* data tweet.

Tabel 1. Hasil *Preprocessing* data

Tahapan	Hasil *Preprocessing*
Data awal	Penggugat di MK Minta Kotak Kosong Dihitung Suara Sah di Pilkada Ini Alasannya! https://t.co/j1tPcpqcak
Case folding	penggugat di mk minta kotak kosong dihitung suara sah di pilkada ini alasannya! https://t.co/j1tpcpqcak
Punctuation dan URL *Removal*	penggugat di mk minta kotak kosong dihitung suara sah di pilkada ini alasannya
Tokenizing	['penggugat', 'di', 'mk', 'minta', 'kotak', 'kosong', 'dihitung', 'suara', 'sah', 'di', 'pilkada', 'ini', 'alasannya']
Normalisasi *Slang Word*	['penggugat', 'di', 'mk', 'minta', 'kotak', 'kosong', 'dihitung', 'suara', 'sah', 'di', 'pilkada', 'ini', 'alasan']
Stopword Removal	['gugat', 'di', 'mk', 'minta', 'kotak', 'kosong', 'hitung', 'suara', 'sah', 'di', 'pilkada', 'ini', 'alas']
Stemming	['gugat', 'mk', 'kotak', 'kosong', 'hitung', 'suara', 'sah', 'pilkada', 'alas']

Pelabelan InSet

Setelah tahap *preprocessing*, proses selanjutnya adalah pelabelan sentimen terhadap tweet. Pelabelan dilakukan secara otomatis menggunakan *Indonesia Sentiment Lexicon* (INSET), yaitu kamus sentimen berbahasa Indonesia yang memuat daftar kata beserta nilai polaritasnya (positif atau negatif). Kata-kata yang tidak terdapat dalam kamus tidak diperhitungkan dalam penentuan skor sentimen. Hasil pelabelan ini menjadi dasar klasifikasi sentimen pada tweet yang dianalisis. Tabel 2 merupakan contoh hasil pelabelan dengan menggunakan InSet.

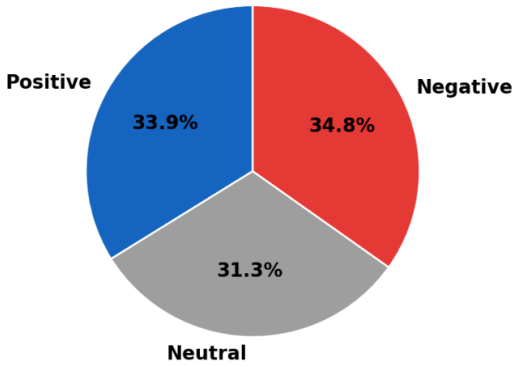
Tabel 2. Contoh Hasil Pelabelan InSet

Tweet	Skor	Sentimen
kpu surabaya terima dukung kotak kosong kursi debat pilkada	2	Positif
gugat mk kotak kosong hitung suara sah pilkada	-4	Negatif
pesta mantap lancar gercos kotak kosong	0	Netral

setiap tweet diklasifikasikan ke dalam kategori positif, negatif, atau netral berdasarkan penjumlahan skor dari kata-kata yang terdapat dalam tweet. Sebagai contoh, tabel 2 menampilkan hasil pelabelan dengan InSet, tweet "kpu surabaya terima dukung kotak kosong kursi debat pilkada", pada kamus InSet, didapatkan skor untuk tiap kata yakni, 'kpu': 0, 'surabaya': 0, 'terima': 2, 'dukung': 4, 'kotak': -3, 'kosong': -4, 'kursi': 0, 'debat': 3, dan 'pilkada': 0. Jika dijumlahkan, total skor sentimen dari tweet ini adalah sebesar dua (2), sehingga dikategorikan sebagai sentimen positif. Selanjutnya, tweet "gugat mk kotak kosong hitung suara sah pilkada", didapatkan skor untuk tiap kata yakni, 'gugat': 0, 'mk': 0, 'kotak': -3, 'kosong': -4, 'hitung': 1, 'suara': 2, 'sah': 0, 'pilkada': 0, jika dijumlahkan maka skornya adalah minus empat (-4), sehingga dikategorikan sebagai sentimen negatif. Tweet "pesta mantap lancar gercos kotak kosong" didapatkan skor untuk tiap kata yakni 'pesta': 4, 'mantap': 5, 'lancar': -2, 'gercos': 0, 'kotak': -3, 'kosong': -4, jika dijumlahkan maka skornya adalah jika dijumlahkan maka skornya adalah nol (0), sehingga dikategorikan sebagai sentimen netral.

Visualisasi Pelabelan Inset

Visualisasi pelabelan dilakukan untuk melihat sebaran sentimen dari tweet yang telah dilabeli menggunakan kamus InSet. *Pie chart* digunakan untuk menggambarkan proporsi tweet dengan sentimen positif, negatif, dan netral, sehingga memberikan gambaran umum kecenderungan opini publik sebelum tahap pelatihan model dilakukan.



Gambar 10. Proporsi Hasil Pelabelan InSet

Berdasarkan visualisasi *pie chart* distribusi sentimen pada Gambar 10, dapat disimpulkan bahwa persepsi masyarakat terhadap isu kotak kosong dalam Pilkada 2024 cenderung seimbang. Dari total 2.168 tweet yang dianalisis, sebanyak 754 tweet (34,8%)

tergolong sentimen negatif, sementara 735 tweet (33,9%) menunjukkan sentimen positif. Adapun 678 tweet lainnya (31,3%) dikategorikan sebagai netral. Meskipun sentimen negatif sedikit lebih dominan, hal ini mencerminkan adanya kecenderungan sebagian masyarakat untuk menyampaikan kritik, kekhawatiran, atau ketidaksetujuan terhadap keberadaan kotak kosong dalam proses pemilihan kepala daerah. Namun demikian, distribusi sentimen secara keseluruhan tidak menunjukkan dominasi pada salah satu kategori, sehingga dapat dikatakan bahwa opini publik relatif seimbang.

GloVe Embedding

Setiap kata pada tweet direpresentasikan dalam bentuk vektor numerik menggunakan metode *GloVe*. Penelitian ini menggunakan *pre-trained GloVe* berdimensi 50 dengan 331.286 kata, yang diambil dari penelitian hasil penelitian terkait klasifikasi perintah dalam aplikasi chatbot [15]. Model ini dipilih karena dilatih dengan menggunakan bahasa Indonesia, sehingga sesuai dengan data yang digunakan dalam penelitian ini. Tabel 3 adalah contoh vektor kata pada *pre-trained GloVe*.

Tabel 3. Contoh Vektor Kata Pada *GloVe*

Kata	Dimensi ke-1	Dimensi ke-2	...	Dimensi ke-50
kotak	0,3744	0,5586	…	-0,6401

Algortima LSTM

Penentuan parameter pelatihan pada model LSTM disesuaikan dengan karakteristik data agar proses pembelajaran berlangsung optimal, serta untuk mencegah terjadinya *overfitting* maupun *underfitting*. Parameter yang digunakan dalam pelatihan model adalah sebagai berikut.

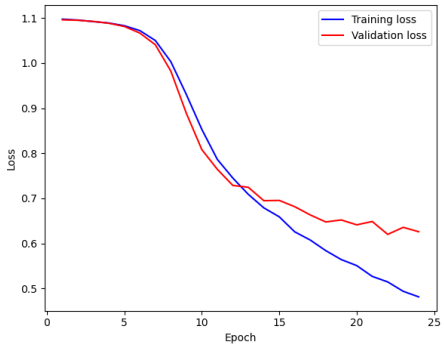
Tabel 4. Parameter Model LSTM

Parameter	Nilai/Keterangan
Dimensi	50
Panjang input	55
Lapisan LSTM	Lapisan pertama 64 unit Lapisan kedua 32 unit
Dropout	0,5
Fungsi aktivasi	*Softmax* (pada lapisan *output*)
Batch size	64
Learning Rate	0,0001
Loss function	*Categorical Crossentropy*
Training Set	90%
Testing set	10%
Jumlah *fold*	10

Tabel 4 menunjukkan parameter yang digunakan dalam membangun model LSTM. Pemilihan parameter dilakukan melalui serangkaian eksperimen untuk menguji berbagai kombinasi nilai, dengan tujuan memperoleh performa model yang paling optimal. Konfigurasi ini dirancang agar model tidak hanya menghasilkan klasifikasi sentimen yang akurat pada data latih, tetapi juga mampu melakukan generalisasi dengan baik terhadap data yang belum pernah dilihat sebelumnya.

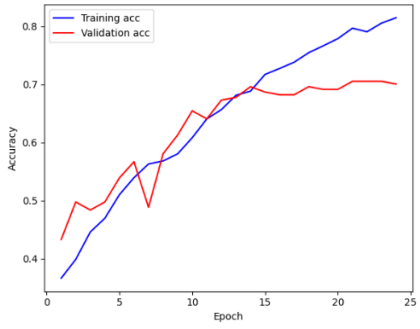
Training Model

Hasil pelatihan model, dapat dilihat melalui Gambar 11 dan 12. Gambar 11 menunjukkan tingkat kesalahan, sedangkan gambar 12 menunjukkan tingkat keberhasilan prediksi model. Kedua plot ini membantu mengidentifikasi adanya *overfitting* atau konvergensi akurasi, serta menentukan waktu yang tepat untuk menghentikan pelatihan guna memperoleh hasil terbaik.



Gambar 11. Plot *Training* dan *Validation Loss*

Dari Gambar 11 dapat dilihat bahwa *training loss* mengalami penurunan yang signifikan dan kemudian menjadi stabil, sementara *validation loss* mulai menunjukkan tanda-tanda stagnasi setelah memasuki *epoch* ke-20. Pola ini mengindikasikan bahwa model mulai mengalami *overfitting*, meskipun pada tahap awal masih menunjukkan kemampuan generalisasi yang baik. Untuk mengatasi hal tersebut dan mencegah penurunan performa pada data validasi, diterapkan teknik *early stopping*.



Gambar 12. Plot *Training* dan *Validation Accuracy*

Gambar 12 menunjukkan adanya tren peningkatan pada *training accuracy*. Sementara itu, *validation accuracy* titik tertinggi pada *epoch* ke-20, sedangkan akurasi pelatihan terus meningkat. Jarak antara *training* dan *validation accuracy* tidak terlalu besar, artinya model masih mampu melakukan generalisasi dengan cukup baik.

Hasil *K-Fold Cross-Validation*

Model GloVe-LSTM dievaluasi menggunakan metode *10-fold cross-validation* untuk mengukur keandalan dan kemampuan generalisasi model. Dataset dibagi menjadi 10 bagian, di mana satu bagian digunakan untuk pengujian dan sisanya untuk pelatihan, secara bergiliran. Metode ini memberikan hasil evaluasi yang lebih menyeluruh dan meminimalkan bias dari pembagian data acak. Tabel 5 berikut menyajikan hasil akurasi, presisi, *recall*, dan *F1-score* dari setiap *fold*.

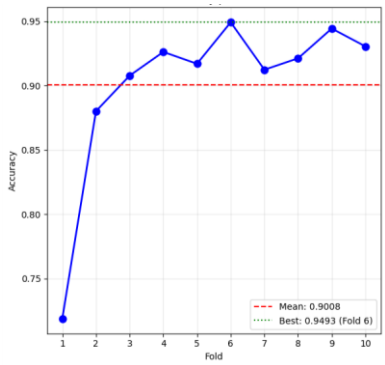
Tabel 5. Metrik Evaluasi *Fold* 10

Fold	Kelas	Akurasi	Presisi	*Recall*	*F1-Score*
1	Positif	71,89%	81,25%	83,87%	82,54%
	Negatif		81,43%	67,06%	73,55%
	Netral		56,63%	67,14%	61,44%
2	Positif	88,02%	97,26%	93,42%	95,30%
	Negatif		80,52%	92,54%	86,11%
	Netral		86,57%	78,38%	82,27%
3	Positif	90,78%	98,61%	93,42%	95,95%
	Negatif		88,75%	92,21%	90,45%
	Netral		84,62%	85,94%	85,27%
4	Positif	92,63%	96,92%	94,03%	95,45%
	Negatif		92,68%	95,00%	93,83%
	Netral		88,57%	88,57%	88,57%
5	Positif	91,71%	100,00%	95,00%	97,44%
	Negatif		87,10%	88,52%	87,80%
	Netral		87,34%	90,79%	89,03%
6	Positif	94,93%	98,53%	94,37%	96,40%
	Negatif		94,12%	97,56%	95,81%
	Netral		92,19%	92,19%	92,19%
7	Positif	91,24%	92,94%	95,18%	94,05%
	Negatif		91,55%	92,86%	92,20%
	Netral		88,52%	84,38%	86,40%
8	Positif	92,13%	98,41%	95,38%	96,88%
	Negatif		94,94%	89,29%	92,02%
	Netral		83,78%	92,54%	87,94%
9	Positif	94,17%	98,48%	94,20%	96,30%
	Negatif		92,59%	97,40%	94,94%
	Netral		91,53%	90,00%	90,76%
10	Positif	93,06%	98,77%	94,12%	96,39%
	Negatif		91,78%	93,06%	92,41%
	Netral		87,10%	91,53%	89,26%
Rata-rata					
	Positif	90,06%	96,12%	93,30%	94,67%
	Negatif		89,55%	90,55%	89,91%
	Netral		84,68%	86,14%	85,31%

Berdasarkan hasil evaluasi *10-fold cross-validation* pada Tabel 5, model *GloVe* dengan LSTM menunjukkan performa yang cukup kuat dalam mengklasifikasikan sentimen terhadap isu *kotak kosong* pada Pilkada 2024. Akurasi model secara umum meningkat dari *fold* pertama hingga *fold* ke-10, dengan tren positif yang mengindikasikan peningkatan kemampuan model dalam mengenali pola sentimen. Kinerja terbaik ditunjukkan pada kelas positif, dengan rata-rata presisi sebesar 96,12%, *recall* 93,30%, dan *f1-score* 94,67%, menandakan model sangat baik dalam mendeteksi tweet bermuatan positif. Untuk sentimen negatif, model mencatat presisi 89,55%, *recall* 90,55%, dan *f1-score* 89,91%, yang menunjukkan ketepatan klasifikasi yang cukup tinggi. Sementara itu, kelas netral mencatat kinerja yang sedikit lebih rendah, dengan *f1-score* rata-rata 85,31%, kemungkinan disebabkan oleh sifat tweet netral yang cenderung ambigu dan kontekstual. Rata-rata akurasi keseluruhan model berada di atas 90%, memperkuat bukti bahwa model memiliki kemampuan generalisasi yang baik dan layak digunakan dalam analisis opini publik berbasis media sosial.

Pemilihan Model Terbaik

Setelah dilakukan evaluasi menyeluruh terhadap model *GloVe* dengan LSTM menggunakan *10-fold cross-validation*, proses selanjutnya yakni mengidentifikasi *fold* dengan kinerja terbaik sebagai representasi model yang paling optimal.



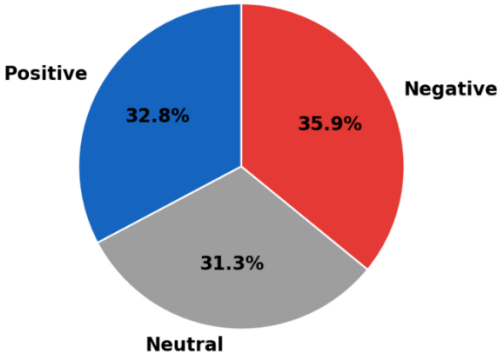
Gambar 13 Plot Training dan Validation Accuracy

Visualisasi pada Gambar 13 di atas menunjukkan akurasi model pada masing-masing *fold*. *Fold* ke-6 menunjukkan performa tertinggi dengan akurasi sebesar 0,9493 atau 94,93%, yang ditandai dengan garis putus-putus berwarna hijau. Sementara itu, rata-rata akurasi dari seluruh *fold* berada pada angka 0,9008 atau 90,08%, ditunjukkan dengan garis putus-putus merah. Berdasarkan pencapaian tersebut, *fold* ke-6 dipilih sebagai model terbaik karena menghasilkan akurasi paling tinggi di antara semua *fold*. Model ini kemudian disimpan dan dijadikan referensi utama untuk melakukan klasifikasi

sentimen terhadap tweet terkait kebijakan *kotak kosong* pada Pilkada 2024.

Visualisasi Hasil Prediksi Sentimen Model

Hasil prediksi model menunjukkan distribusi sentimen masyarakat ke dalam tiga kelas. Distribusi ini divisualisasikan dalam bentuk *pie chart* untuk memudahkan pemahaman proporsi masing-masing sentimen dalam dataset penelitian.



Gambar 14 Visualisasi Proporsi Sentimen Hasil Prediksi Model GloVe LSTM

pada Gambar 14, hasil prediksi sentimen dari model terbaik *GloVe* dengan LSTM pada *fold* ke-6 menunjukkan distribusi sebagai berikut: 35,9% tweet diklasifikasikan sebagai negatif, 32,8% sebagai positif, dan 31,3% sebagai netral. Komposisi ini mengindikasikan adanya kecenderungan opini publik yang bernuansa negatif terhadap kebijakan *kotak kosong* dalam Pilkada 2024. Dominasi sentimen negatif tersebut dapat diartikan sebagai respons masyarakat yang berupa kritik, kekhawatiran, atau bentuk penolakan terhadap keberadaan kebijakan tersebut dalam proses pemilihan kepala daerah.

Visualisasi *Word Cloud*

Tweet yang telah diklasifikasikan oleh model LSTM, kemudian divisualisasikan dalam bentuk *wordcloud* untuk tiap kelas sentimennya. Visualisasi ini bertujuan untuk menampilkan kata-kata yang paling sering muncul dalam dari tiap kelas sentimen..

Gambar 15 Wordcloud Kelas Positif Pada Kotak Kosong Pilkada 2024

Wordcloud untuk sentimen positif pada gambar 15 menampilkan kata-kata utama yang mencerminkan sikap penerimaan dan dukungan terhadap pelaksanaan Pilkada dengan keberadaan *kotak kosong*. Istilah "kotak kosong" menggambarkan pandangan masyarakat bahwa proses Pilkada tetap dapat berlangsung dengan baik dan sukses tanpa hambatan berarti. Kata "pilkada" sendiri muncul sebagai simbol proses demokrasi yang berjalan lancar dan tetap sah meskipun dihadapkan pada pilihan *kotak kosong*. Kata "lawan" dalam konteks ini merujuk pada keberhasilan menghadapi calon tunggal, sedangkan kata "menang" mencerminkan optimisme dan dukungan terhadap kemenangan yang diraih oleh pilihan *kotak kosong*. Secara keseluruhan, kata-kata ini menunjukkan adanya pandangan positif dan harapan terhadap keberadaan *kotak kosong* dalam Pilkada 2024. Hasilnya terdapat pada Gambar 16.

Gambar 16 Wordcloud Kelas Negatif Pada Kotak Kosong Pilkada 2024

Wordcloud sentimen negatif pada gambar 16 memperlihatkan kata-kata yang merefleksikan kritik publik terhadap kebijakan *kotak kosong* dalam Pilkada 2024. Istilah "kotak kosong" yang paling menonjol sering digunakan dalam konteks kekecewaan masyarakat atas terbatasnya pilihan calon yang layak. Kata "kalah" mengindikasikan kekhawatiran terhadap kemungkinan kekalahan calon petahana ketika berhadapan dengan *kotak kosong*. Sementara itu, kata "tidak" mencerminkan ekspresi ketidaksetujuan, baik terhadap partai politik tertentu yang dianggap bertanggung jawab atas munculnya calon tunggal, maupun terhadap ketidakpercayaan terhadap opsi lain yang tersedia. Kata "partai" sendiri menunjukkan kritik langsung terhadap dominasi partai politik dalam proses Pilkada, serta harapan masyarakat akan munculnya calon independen sebagai alternatif.

Gambar 17 Wordcloud Kelas Netral Pada Kotak Kosong Pilkada 2024

Wordcloud untuk sentimen netral pada Gambar 17 menampilkan kata yang menggambarkan diskusi informatif mengenai Pilkada dengan keberadaan *kotak kosong*. Kata "kotak kosong" kerap muncul dalam konteks proses Pilkada tetap berlangsun. Istilah "pilkada" mencerminkan pembahasan seputar pelaksanaan, partisipasi pemilih, dan hasil pemilu tanpa memihak pada salah satu sudut pandang, baik mendukung maupun menolak *kotak kosong*. Kata "paslon" (pasangan calon) menunjukkan perhatian terhadap daerah-daerah yang menyelenggarakan Pilkada dengan konfigurasi kotak kosong versus calon tunggal, sebagai bagian dari laporan situasi. Sementara itu, kata "calon tunggal" mengindikasikan kondisi di mana hanya terdapat satu kandidat yang secara resmi maju dalam kontestasi politik tersebut.

4. KESIMPULAN

Penelitian ini membuktikan bahwa model *GloVe* dengan LSTM berhasil mengklasifikasikan sentimen masyarakat terhadap kebijakan *kotak kosong* pada Pilkada 2024 berdasarkan data Twitter. Dari 2.168 tweet yang dianalisis, distribusi sentimen menunjukkan 35,9% negatif, 32,8% positif, dan 31,3% netral. Hal ini menunjukkan opini publik yang cukup seimbang, dengan proporsi sedikit lebih banyak pada sentimen negatif, yang mengindikasikan ketidakpuasan terhadap sistem politik yang memungkinkan calon tunggal. *Wordcloud* sentimen menunjukkan bahwa sentimen positif didominasi oleh dukungan terhadap kelancaran Pilkada dan optimisme terhadap kotak kosong, sementara sentimen negatif berisi kritik terhadap dominasi partai dan keterbatasan pilihan politik. Sentimen netral umumnya bersifat informatif. Dari segi performa, model *GloVe*-LSTM mencatat rata-rata akurasi 90,06% pada skema *10-fold cross-validation*, dengan akurasi tertinggi sebesar 94,93% pada *fold* ke-6. Kinerja per kelas menunjukkan bahwa sentimen positif diklasifikasikan paling akurat (*F1-score* 94,67%), diikuti oleh negatif (*F1-score* 89,91%), dan netral (*F1-score* 85,31%). Nilai yang lebih rendah pada kelas netral menunjukkan tantangan model dalam mengidentifikasi tweet yang bersifat ambigu. Performa yang mulai stabil pada *fold* ke-6 hingga ke-10 menandakan kemampuan model untuk melakukan generalisasi setelah proses pelatihan yang optimal. Secara keseluruhan, model LSTM berbasis *GloVe* terbukti efektif untuk klasifikasi sentimen terkait isu sosial politik di media sosial, dan berpotensi digunakan sebagai dasar dalam pengembangan sistem pemantauan opini publik secara otomatis. Meskipun penelitian ini menunjukkan hasil yang baik, terdapat beberapa keterbatasan yang perlu diketahui. Pada tahap *preprocessing* masih ditemukan kata berbahasa asing yang tidak terfilter, sehingga pada penelitian selanjutnya disarankan untuk menambahkan proses terjemahan agar data dapat lebih konsisten dalam bahasa Indonesia. Selanjutnya, banyak kata *slang* yang belum terindeks dalam kamus InSet yang ada pada sumber daring yang tersedia, sehingga diperlukan pemantauan dan pengayaan kamus secara berkelanjutan agar hasil *preprocessing* lebih sesuai dengan kebutuhan analisis sentimen. Pelabelan sentimen yang dilakukan secara otomatis menggunakan *Indonesia Sentiment Lexicon (INSET)* memiliki keterbatasan dalam menangkap konteks yang lebih kompleks, seperti ironi, sarkasme, atau ambiguitas bahasa di media sosial. Selain itu, penelitian ini hanya menggunakan satu pendekatan utama, yaitu kombinasi *GloVe* dan *LSTM*, sehingga belum ada perbandingan dengan model lain seperti *Bidirectional LSTM (BiLSTM)*, *GRU*, atau model berbasis *transformer* seperti *BERT* dan *IndoBERT*.

DAFTAR PUSTAKA

[1] S. Sarbaini, "Penyelenggaraan Pemilihan Kepala Daerah Secara Langsung dan Demokratis Sebagai Bentuk Perwujudan Hak Asasi Politik Masyarakat di Indonesia," *Legalitas: Jurnal Hukum*, vol. 12, no. 1, p. 107, Jun. 2020, doi: 10.33087/legalitas.v12i1.197.

[2] A. Budiman, "PILKADA PASLON TUNGGAL, KINERJA PARTAI POLITIK, DAN MASA DEPAN DEMOKRASI DI INDONESIA," 2018.

[3] Simon Kemp, "Digital 2023: Indonesia," www.datareportal.com.

[4] M. Athoillah and R. K. Putri, "Utilizing Support Vector Machines to Detect Hate Speech on Social Media," *Science, Engineering and Technology*, vol. 4, no. 2, pp. 53–60, Sep. 2024, doi: 10.54327/set2024/v4.i2.129.

[5] R. Kusumaningrum, I. Z. Nisa, R. P. Nawangsari, and A. Wibowo, "Sentiment analysis of Indonesian hotel reviews: from classical machine learning to deep learning," *International Journal of Advances in Intelligent Informatics*, vol. 7, no. 3, pp. 292–303, Nov. 2021, doi: 10.26555/ijain.v7i3.737.

[6] B. Jang, M. Kim, G. Harerimana, S. U. Kang, and J. W. Kim, "Bi-LSTM model to increase accuracy in text classification: Combining word2vec CNN and attention mechanism," *Applied Sciences (Switzerland)*, vol. 10, no. 17, Sep. 2020, doi: 10.3390/app10175841.

[7] Kusum and S. P. Panda, "Sentiment analysis using global vector and long short-term memory," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no.

1, pp. 414–422, Apr. 2022, doi: 10.11591/ijeecs.v26.i1.pp414-422.

[8] A. Rao and N. Spasojevic, "Actionable and Political Text Classification using Word Embeddings and LSTM," Jul. 2016, [Online]. Available: http://arxiv.org/abs/1607.02501

[9] Xianghua Fu, Jingying Yang, Jianqiang Li, Min Fang, and Huihui Wang, "Lexicon-Enhanced LSTM With Attention for General Sentiment Analysis," *IEEE*, vol. 6, Oct. 2018.

[10] A. Nilsen, "Perbandingan Model RNN, Model LSTM, dan Model GRU dalam Memprediksi Harga Saham-Saham LQ45," *Jurnal Statistika dan Aplikasinya*, vol. 6, no. 1, 2022.

[11] Safwan Mahmood Al-Selwi *et al.*, "RNN-LSTM: From applications to modeling techniques and beyond—Systematic review," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 5, Jun. 2024.

[12] X. Wu, B. Xiang, H. Lu, C. Li, X. Huang, and W. Huang, "Optimizing Recurrent Neural Networks: A Study on Gradient Normalization of Weights for Enhanced Training Efficiency," *Applied Sciences (Switzerland)*, vol. 14, no. 15, Aug. 2024, doi: 10.3390/app14156578.

[13] Palash Goyal, Sumit Pandey, and Karan Jain, *Deep Learning for Natural Language Processing: Creating Neural Networks with Python*. Bangalore: apress, 2018.

[14] A. A. Winoto and A. F. V. Roy, "Model of Predicting the Rating of Bridge Conditions in Indonesia with Regression and K-Fold Cross Validation," *International Journal of Sustainable Construction Engineering and Technology*, vol. 14, no. 1, pp. 249–259, Feb. 2023, doi: 10.30880/ijscet.2023.14.01.022.

[15] D. S. Soper, "Greed is good: Rapid hyperparameter optimization and model selection using greedy k-fold cross validation," *Electronics (Switzerland)*, vol. 10, no. 16, Aug. 2021, doi: 10.3390/electronics10161973.

[16] IRFAN HANIF, "KLASIFIKASI PERINTAH BAHASA NATURAL MENGGUNAKAN GLOBAL VECTORS FOR WORD REPRESENTATIONS (GLOVE), CONVOLUTIONAL NEURAL NETWORKS, DAN TEKNIK TRANSFER LEARNING PADA APLIKASI CHATBOTS," Institut Teknologi Sepuluh Nopember, Surabaya, 2018.