

ANALISIS SENTIMEN UJARAN KEBENCIAN PADA TWEET DI TWITTER

Talita Cintiana Adisti , Erna Daniati, Aidina Ristiyawan

Sistem Informasi, Universitas Nusantara PGRI Kediri

Jl.KH.Ahmad Dahlan No.76, Mojoroto, Kediri 64112

talitacintianaadisti244@gmail.com

ABSTRAK

Media Sosial merupakan platform utama bagi informasi, tetapi juga sensitif terhadap penyebaran ujaran kebencian yang mempengaruhi individu maupun masyarakat secara negatif. Penelitian ini memiliki tujuan untuk menganalisis sentimen ujaran kebencian terhadap tweet di twitter, Menggunakan pengolahan bahasa alami atau (Natural Language Processing/NLP) Metodologi di gunakan dalam pemrosesan data preprocessing yaitu seperti case foldig, tokenisasi, stopword removal. Selanjutnya Support Vector Machine atau disebut (SVM) menunjukkan bahwa model algoritma tertentu memiliki wawasan mengenai pola penyebaran pada ujaran kebencian di Twitter, dapat digunakan sebagai dasar pengembangan sistem deteksi hasil ini diharapkan mendukung ruang di gital yang lebih aman dan inklusif

Kata kunci : Analisis Sentimen Ujaran Kebencian, Twitter, Support Vector Machie

1. PENDAHULUAN

Saat ini penggunaan media sosial dapat dikatakan sebagai salah satu bentuk kegiatan yang tidak bisa dilepaskan dari keseharian kita, media sosial dapat dijadikan sebagai sebuah sarana untuk melakukan komunikasi, berbagai informasi dan juga untuk mengetahui suatu pendapat atau dampaknya. Dari penggunaan media sosial dapat memberikan hal yang bersifat positif maupun negatif yang dapat mempengaruhi kehidupan seseorang.

[1] Twitter menjadi platform kelima yang paling banyak dan aktif dalam penggunaan Twitter. Kata yang disampaikan dalam Twitter terkadang tidak dituliskan dalam bentuk terstruktur. Dengan Tingginya tingkat ujaran kebencian dapat mengganggu kesejahteraan dan emosional seseorang. Sebagai contoh kita bisa melihat respon masyarakat dalam menanggapi suatu kejadian yang sedang viral diberitakan, hal ini dapat digunakan untuk menilai sentimen masyarakat terhadap sebuah topik. Kemudahan bagi seseorang untuk menyampaikan pendapat kini dapat diwujudkan melalui media sosial sehingga diterbitkan Undang-Undang Informasi dan Transaksi Elektronik (UU ITE) dikutip (Oryza Habibie Rahman et al., 2021)

[2] Untuk membatasi ruang bagi masyarakat dalam mengutarakan pendapat di dunia maya. Salah satu permasalahan yang masuk ke dalam UU ITE adalah mengenai ujaran kebencian. Klasifikasi teks adalah bagian dari text mining, merupakan suatu metode untuk mempresikisi kategori kelas dari sebuah data. Beberapa contoh metode klasifikasi yang dapat digunakan untuk digunakan melakukan klasifikasi teks adalah metode K-Nearst Neighbor (KKN), Support Vector Machine (SVM). Namin yang paling umum sering digunakan untuk. [3] klasifikasi teks adalah SVM. Performa dari algoritma yang digunakan untuk klasifikasi teks tersebut sangat dipemilihan fitur terdapat teknik untuk melakukan ekstraksi fitur seperti N-gram Inverse Document

2. TINJAUAN PUSTAKA

2.1. Kajian Teori

Dari Teori Penelitian (Hilmi Zain et al., n.d) Penelitian ini menggunakan pendekatan kinerja tiga algoritma klasifikasi SVM, KNN, dan Naïve Bayes dalam analisis sentimen Terdapat data Twitter terkait calon preside 2024. Sama seperti penelitian Saya sama sama menggunakan SVM. Tetapi perbedaya pada data saya membaca aspek adanya ujaran kebencian

2.2. Tabel Penelitian Terdahulu

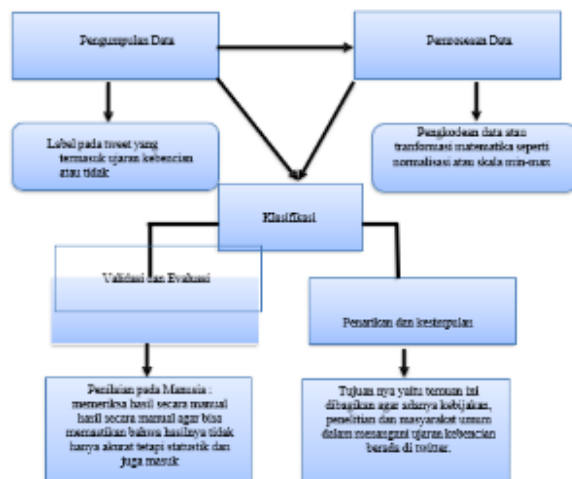
Pengarangnya Al & Mubarak judulnya Analisis sentimen penggunaan Twitter Terhadap kebijakan pemberlakuan sosial berskala besar (PSBB) menggunakan metode naïve Bayes Hasilnya yaitu Proses Klasifikasi menggunakan perangkat lunak. Penelitian ini tentang dataset digunakan memiliki 64 tweet negatif sedangkan tweet netral 60 dan 34 tweet positif sehingga total tweet keseluruhan sebanyak 158 tweet

Pengarangnya Pangestuti berjudul Klasifikasi komentar Abusive Hate Speech Teks Twitter Menggunakan Metode Convolution Neural Network Hasilnya yaitu Proses optimal training menggunakan akurasi sebagai metrik pengukuran validasiya optimizer Adam, dan loss validasi menggunakan history crpss entropy untuk mendapatkan model. Training terpisah anantara kelas HS

Pengarangnya Adhari, judulnya Identifikasi Ujaran Kebencian di Twitter menggunakan Metode Convolutional Neural Network Hasilnya yaitu menggunakan dataset yang berjumlah 2.346 data tweet hasil dari dari pengumpulan data

2.3. Kerangka Berpikir

Kerangka Berpikir meliputi Pengumpulan data, Pemrosesan Data, Klasifikasi, Validasi dan Evaluasi, Penarik dan Kesimpulan



Gambar 1. Kerangka Berfikir

2.4. Pengumpulan Data : Tahap pengumpulan data

Untuk analisis sentimen Ujaran kebencian pada Tweet Twitter dilakukan dalam aplikasi pengumpulan data Twitter menggunakan alat-alat khusus mengambil Data dari platform tersebut.

2.5. Pemrosesan Data atau (Stopwords Removal)

Pemrosesan Data dalam Konteks meliputi proses menghilangkan kata-kata umum tidak Memiliki makna signifikan dalam analisis sentimen.

2.6. Klasifikasi atau (Pembersih Data)

Proses Klasifikasi, data Tweet perlu dibersihkan dan berbagai yang tidak relevan dan Mengganggu akurasi Model.

2.7. Validasi dan Evaluasi

Validasi yaitu model dalam membagi Dataset menjadi beberapa subset (menggunakan k-fold cross-validation).

2.8. Penarikan dan Kesimpulan

Proses dan validasi dalam tahap kesimpulan dilakukan. Yaitu Hasil analisis sentimen ujaran kebencian pada Tweet di Twitter memberi pemahaman lebih baik tentang sentimen publik dan konten negatif di media sosial. Untuk kesimpulan dapat memberikan wawasan berharga bagi peneliti membuat kebijaksanaan pada platform media sosial dalam mengatasi masalah pada ujaran kebencian.

2.9. Hasil dan Pembahasan

Hasil penelitian menunjukkan menggunakan metode yang digunakan juga memiliki ruang peningkatam. Pembahasan ini penting pada penelitian lebih lanjut untuk meningkatkan akurasi dan kendala deteksi ujaran kebencian

3. METODE PENELITIAN

3.1. Metode Penelitian

Metode Peneliian yang digunakan dalam penelitian ini menggunakan mode kualitatif karena melibatkan pengumpulan data yang dilakukan dengan cara observasi,wawamcara dan dokumentasi.

3.2. Teknik Penelitian

Konteks ujaran kebencian diteliti dapat memberikan wawasan yang lebih baik. Misalnya jika penelitian dilakukan menganalisis ujaran kebencian dalam konteks politik suatu negara Penelitian dapat dilakukan di negara tersebut Penelitian ini dilakukan dimana saja dengan adanya akses internet memadai. Karena Twitter adalah platform daring. Penelitian dapat mengakses data tweet dari mana saja di dunia. Namun lokasi penelitian yang relevan dengan konteks ujaran kebencian yang diteliti.

Sedangkan teknik penelitian yang digunakan adalah eksperimen karena diperlukan mengumpulkan data lalu pemrosesan data dan klasifikasi maupun validasi dan Evaluasi.

3.3. Objek dan Waktu penelitian

Objek yang diteliti adalah tentang cuitan masyarakat di Twitter tentang hate Speech atau ujaran kebencian. Yang tersedia secara publik dengan URL: (ID-Hatespeech-Deteciob-Master,n.d).

Tabel 1. Waktu Penelitian

no	Kegiatan	Waktu Penelitian																							
		Juni 2024			Juli 2024			Agustus 2024			September 2024			Oktober 2024			November 2024			Desember 2024			Januari 2025		
		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
1	Waktu Penelitian																								
2	Mengumpulkan data																								
3	Pemrosesan data																								
4	Analisa Sentimen																								
5	Klasifikasi Ujaran Kebencian																								
6	Evaluasi & Validasi																								
7	Interpretasi & Visualisasi																								

3.4. Populasi

Data bernama hate Speech tentang cuitan komentar di Twitter megggunakan fitur Machine Learning yaitu pembelajaran mesin yang dapat digunakan melatih miodel yang dapat mengklasifikasikan twettt menjadi positif negatif atau netral Dalam hal ini data diperoleh 10 % data dalam populasi.

4. HASIL DAN PEMBAHASAN

4.1. Import Libraries (Library Python)



Gambar 1. Import Libraries (Library Python)

Import Libraries adalah tentang import digunakan untuk analisis data dan visualisasi ada Import Pandas : yaitu untuk mengelola dan menganalisis data dalam bentuk tabel seperti Excel atau SQL. Import numpy as np : yaitu digunakan untuk operasi matematika dan array (seperti matriks) Import matplotlib.pyplot ada di Matplotlib pyplot : membuat grafik atau visualisasi seperti garis, batang dan lingkaran sedangkan Import seaborn as sns : adalah visualisasi data lebih estetik dan mudah digunakan daripada Matplotlib % Matplotlib inline : jupyter notebook membuat kode dijalankan. Sns.set(style='darkgrid') Mengatur gaya tampilan ke dalam latar gelap dan grid karena mempermudah membaca pada data Import warnings : Pustaka mengatur peringatan pada Python Warnings. fileterwarning('ignore') Mengabaikan pesan peringatan dari output karena lebih bersih.

4.2. Cloud Data

Data terlihat

```
[2]: tweet_df = pd.read_csv('dataset/hate_speech.csv', names=['sentimen', 'tweet'])
tweet_df = pd.DataFrame(tweet_df)
```

```
[3]: tweet_df
```

	sentimen	tweet
0	Label	
1	Non_HS	RT @spardayz: Fadli Zon Minta Mendagri Segera...
2	Non_HS	RT @baguscondromoc: Mereka terus melukai aksi...
3	Non_HS	Sybil bagaimana gubernur melakukan kekerasan...
4	Non_HS	Ahmad Dhani Tak Paus Debat Pilkada, Masalah Ja...
...	...	...
709	HS	Muka Si BABI Ahok Tuh Yg Mirip SERBET Lantai...
710	HS	Betul bang hancuran merk bang munahkan chi...
711	HS	Sapa yg bilang Ahok anti korupsi? klo grombo...
712	HS	Gw juga ngimpi SENTILIN BUI BABI AHOK postar...
713	HS	Mudahkan gw ketemu sama Si BABI IWAN BOPENG DI...

714 rows x 2 columns

Gambar 2. Cloud Data

Cloud Data adalah data tweet yang sudah diklasifikasikan menjadi kalimat ujaran kebencian ataupun tidak. Dataset berisi dari kumpulan tweet yang sudah di beri label sentimen. Digunakan untuk mengidentifikasi tweet yang dimana Analisis tersebut mengandung ujaran kebencian ('HS') ataupun ('Non_HS') atau tidak. Jumlah data sekisar 714 baris yang berarti ada 714 tweet dalam dataset. Hasil dari file CSV disimpan dalam objek DataFrame diberi nama tweet_df berarti struktur data saat digunakan untuk menyimpan data

4.3. Populasi dan Sampel

Populasi dan Sample

```
[2]: import random as rd
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')

# Membaca dataset
file_path = 'dataset/hate_speech.csv'
tweet_df = pd.read_csv(file_path, names=['sentimen', 'tweet'])

# Menampilkan 5 data teratas
print('Data Teratas:')
print(tweet_df.head())

# Informasi dataset
print('Informasi Dataset:')
print(tweet_df.info())

# Mengacak distribusi label sentimen
print('Distribusi Sentimen:')
print(tweet_df['sentimen'].value_counts())

# Menghapus data (menghapus duplikat)
tweet_df_cleaned = tweet_df.drop_duplicates(subset='tweet', keep='first')

# Menghapus data dengan nilai null
tweet_df_cleaned = tweet_df_cleaned.dropna()

# Menampilkan data setelah dibersihkan
print('Data Setelah Dibersihkan:')
print(tweet_df_cleaned.info())

# Statistik deskriptif dataset
print('Statistik Deskriptif:')
print(tweet_df_cleaned.describe())

# Menyimpan data yang sudah dibersihkan ke file baru
cleaned_file_path = 'dataset/hate_speech_cleaned.csv'
tweet_df_cleaned.to_csv(cleaned_file_path, index=False)
print('Data yang telah dibersihkan disimpan di', cleaned_file_path)
```

Gambar 3. Populasi dan Sampel

Populasi adalah kumpulan data dari objek yang teliti Populasi juga merupakan dari semua tweet yang ada pada hate_speech.csv mendapatkan 10% karena Bagian dari data dalam populasi di acak oleh sampel mengambil 10 %

4.4. Pengumpulan Data

Mengumpulkan data

```
import pandas as pd

# Membaca dataset dan memberikan nama Anlam
file_path = 'dataset/hate_speech.csv'
tweet_df = pd.read_csv(file_path, names=['sentimen', 'tweet'])

# Menampilkan 5 data teratas
print('Data Teratas:')
print(tweet_df.head())

# Informasi dataset
print('Informasi Dataset:')
print(tweet_df.info())

# Mengacak distribusi label sentimen
print('Distribusi Sentimen:')
print(tweet_df['sentimen'].value_counts())

# Menghapus data (menghapus duplikat)
tweet_df_cleaned = tweet_df.drop_duplicates(subset='tweet', keep='first')

# Menghapus data dengan nilai null
tweet_df_cleaned = tweet_df_cleaned.dropna()

# Menampilkan data setelah dibersihkan
print('Data Setelah Dibersihkan:')
print(tweet_df_cleaned.info())

# Statistik deskriptif dataset
print('Statistik Deskriptif:')
print(tweet_df_cleaned.describe())

# Menyimpan data yang sudah dibersihkan ke file baru
cleaned_file_path = 'dataset/hate_speech_cleaned.csv'
tweet_df_cleaned.to_csv(cleaned_file_path, index=False)
print('Data yang telah dibersihkan disimpan di', cleaned_file_path)
```

Gambar 4. Pengumpulan Data

Menyimpan DataFrame yang dibersihkan di file CSV menggunakan nama baru bernama hate_speech_cleaned.csv lalu di ubah di dalam dataset. Index flase agar mencegah teks DataFrame terimpan di file CSV Secara Keseluruhan kode ini melakukan pra-pemrosesan data penting karena dapat digunakan untuk analisis model seperti memberishkan data dari duplikat atau nilai yang hilang. Analisis sentimen mengklasifikasikan tweet positif maupun negatif pada pendeteksi tweet yang mengandung ujaran kebencian berada pada topik topik utama juga visualisasi untuk mengeksplorasi data secara mendalam

4.5. Hasil Data

```
0 Label Tweet
1 Non_HS RT @spardayz: Fadli Zon Minta Mendagri Segera...
2 Non_HS RT @baguscondromoc: Mereka terus melukai aksi...
3 Non_HS Sybil bagaimana gubernur melakukan kekerasan...
4 Non_HS Ahmad Dhani Tak Paus Debat Pilkada, Masalah Ja...

Informasi Dataset:
Name: Pandas.core.frame.DataFrame
RangeIndex: 714 entries, 0 to 713
Data columns (total 2 columns):
# Column Non-Null Count Dtype
---
0 Label 714 non-null object
1 tweet 714 non-null object
dtypes: object(2)
memory usage: 11.3+ KB
None

Distribusi Sentimen:
sentimen
Non_HS 473
HS 240
Label 3
Name: count, dtype: int64

Data Setelah Dibersihkan:
Name: Pandas.core.frame.DataFrame
RangeIndex: 714 entries, 0 to 713
Data columns (total 2 columns):
# Column Non-Null Count Dtype
---
0 Label 714 non-null object
1 tweet 714 non-null object
dtypes: object(2)
memory usage: 11.3+ KB
None

Statistik Deskriptif:
count 714
unique 3
top Non_HS
freq 3

Data yang telah dibersihkan disimpan di dataset/hate_speech_cleaned.csv
```

Gambar 5. Hasil data

Hasil ini menunjukkan bahwa Python berhasil membaca dan membersihkan penyimpanan dataset ujaran kebencian. Adanya Penambahan data setelah di bersihkan kemungkinan penggabungan data atau pengisian data yang hilang. Meskipun tidak secara eksplisit telah dilakukan penghapusan duplikat dari kolom tweet bertujuan memastikan hanya muncul satu kali dalam dataset. Juga dilakukan baris yang memiliki tabel sentimen pada tweet tujuannya menghindari data yang tidak lengkap.

4.6. Pemrosesan Data

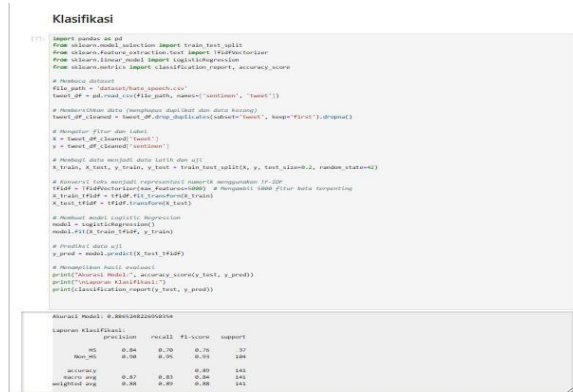
Pemrosesan data adalah Data dibaca, langsung melibatkan membersihkan data (preprocessing), ekstra fitur dan memberikan pelatihan dan evaluasi model klasifikasi. Tweet yang mengandung ujaran kebencian berdasarkan kata kata atau frasa yang dapat kita

gunakan adanya pun visualisasi agar mempermudah pemahaman distribusi data seperti word cloud atau grafik batang.



Gambar 6. Pemrosesan Data

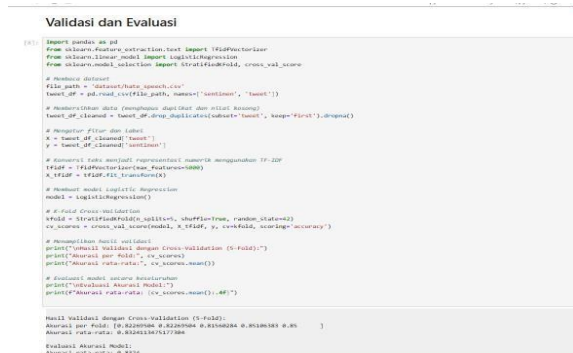
4.7. Klasifikasi



Gambar 7. Klasifikasi

Klasifikasi ujaran kebencian menggunakan Regresi Logistik dan TF-IDF. Langkahnya meliputi pemrosesan data, ekstraksi fitur, pelatihan model, prediksi, dan evaluasi. Hasilnya menunjukkan model. Bahwa performa pada dataset dan parameter yang digunakan.

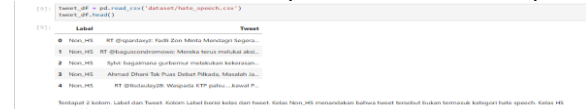
4.8. Validasi dan Evaluasi



Gambar 8. Validasi dan Evaluasi

Melakukan Validasi dan evaluasi untuk model klasifikasi teks. Untuk menghasilkan tweet sebagai “hatespeech (ujaran kebencian) atau bukan dan ada Hyperparameter tuning merupakan penyetelan pada model regresi logistik dan matriks evaluasi adalah akurasi dan metrik yang mempunyai recall, F1- Score terutama kelas data yang tidak seimbang dengan menggunakan visualisasi mendapatkan grafik atau plot yang memiliki pemahaman yang baik

4.9. Tweet DataFrame (Tweet Struktur data)



Gambar 9. Tweet dataframe(Tweet Struktur Data)

File berisi data berupa tweet yang sudah diklasifikasi data tersebut digunakan untuk membangun model klasifikasi. Agar dapat membedakan antara tweet mengandung hate speech (HS) dan yang tidak (Non_HS). Untuk mengidentifikasi hate speech dalam teks. Kode pd.read CSV dapat menyimpan sebuah DataFrame (tabel data) diberi nama tweet sedangkan tweet.head() digunakan untuk menampilkan 5 baris pertama dari Data

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian. Ujaran kebencian di media sosial menjadi masalah kompleks yang membutuhkan penanganan multidimensional. Kombinasi antara regulasi tepat pengembangan teknologi dan peningkatan kesadaran masyarakat adalah kunci menciptakan ruang digital lebih aman dan inklusif dan saran nya. Perlu dilakukan lebih mendalam penelitian ini karena masalah ujaran kebencian di media sosial dan mencari solusi yang komprehensif serta perlu adanya tingkatan sejak dini untuk membekali masyarakat dengan ketrampilan berpikir kritis dan literasi di

DAFTAR PUSTAKA

- [1] Maulana Baihaqi, W., Pinilih, M., & Rohmah, M. (2020). *KOMBINASI K-MEANS DAN SUPPORT VECTOR MACHINE (SVM) UNTUK MEMPREDIKSI UNSUR SARA PADA TWEET*. 7(3). <https://doi.org/10.25126/jtiik.202072126>
- [2] Hilmi Zain, H., Maulana Awannga, R., & Isti Rahayu, W. (n.d.). Perbandingan Model Svm, Knn Dan Naïve Bayes Untuk Analisis Sentiment Pada Data Twitter: Studi Kasus Calon Presiden 2024. *JIMPS: Jurnal Ilmiah Mahasiswa Pendidikan Sejarah*, 8(3), 2083–2093. <https://doi.org/10.24815/jimps.v8i3.25342>
- [3] Akmalia, R., Slamet, I., & Pratiwi, H. (n.d.). Analisis Sentimen Twitter Berbahasa Indonesia Terhadap Aplikasi PeduliLindungi dengan Algoritma SVM, KNN, dan Regresi Logistik. In *Prosiding Seminar Nasional MIPA UNIPA* (Vol. 2022).
- [4] Al, F., & Mubarak, R. (n.d.). *ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP KEBIJAKAN PEMBERLAKUAN PEMBATAAN SOSIAL BERSKALA BESAR (PSBB) DENGAN METODE NAÏVE BAYES*. 7(1),
- [5] Irma Purnamasari, A., & Ali, I. (2024). *ANALISIS SENTIMEN KOMENTAR BERITA DETIK.COM MENGGUNAKAN*

- ALGORITMA SUPORT VEKTOR MACHINE (SVM). In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Issue 3).
- [6] Murni, M., Riadi, I., & Fadlil, A. (2023). Analisis Sentimen HateSpeech pada Pengguna Layanan Twitter dengan Metode Naïve Bayes Classifier (NBC). *JURIKOM (Jurnal Riset Komputer)*, 1566. <https://doi.org/10.30865/jurikom.v10i2.5984>
- [7] Oryza Habibie Rahman, Gunawan Abdillah, & Agus Komarudin. (2021). Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(1), 17–23. <https://doi.org/10.29207/resti.v5i1.2700>
- [8] Pangestuti, I., Agustian, S., Informatika, T., Sains, F., Sultan Syarif Kasim Riau, U., Soebrantas No, J. H., & Baru, S. (2022). *Klasifikasi Komentar Abusive Dan Hate Speech Teks Twitter Menggunakan Metode Convolutional Neural Network*. 7, 2022. <https://radimrehurek.com/gensim/models/word2vec.html>
- [9] Sadat, A., Lawelai, H., & Suherman,) Ansar. (n.d.). *ANALISIS SENTIMEN MEDIA SOSIAL: HATE SPEECH KEPADA PEMERINTAH DI TWITTER*. 10, 2022.
- [10] Watori, J., Aryanti, R., & Junaidi, A. (2020). Penggunaan Algoritma Klasifikasi Terhadap Analisa Sentimen Pemindahan Ibukota Dengan Pelabelan Otomatis. *JURNAL INFORMATIKA*, 7(1). <http://ejournal.bsi.ac.id/ejurnal/index.php/ji>