UNP JOURNAL OF STATISTICS AND DATA SCIENCE
Volume 1 Nomor 5 Tahun 2023 Halaman 385-391.
ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Sentiment Analysis of Prabowo Subianto as 2024 Presidential Candidate on Twitter Using K-Nearest Neighbor Algorithm

Aurumnisva Faturrahmi, Zamahsary Martha*, Yenni Kurniawati, Fadhilah Fitri
Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia
*Corresponding author: zamahsarymartha@fmipa.unp.ac.id

Submitted : 13 September 2023
Revised : 27 Oktober 2023
Accepted : 13 November 2023

ABSTRACT

The presidential election is one of the most talked topics at this moment. Based on many surveys, Prabowo Subianto is one of the strongest candidates for the upcoming 2024 presidential election. This research aims to see how the public sentiment towards Prabowo Subianto as the presidential candidate tends to be positive or negative. Sentiment classification was conducted using the K-Nearest Neighbor (KNN) algorithm. This algorithm classifies sentiment based on the k value of the nearest neighbor. This analysis was conducted in several stages such as data collection, text preprocessing, data labelling, data classification using the KNN algorithm, and evaluating the accuracy of the model in classifying sentiment. In this research, the results of the sentiment classification were 2731 positive sentiments and 76 negative sentiments. Where the accuracy rate produced by the model using the value of k = 3 on the division of training data and testing data of 80:20 is 97,33%.

Keywords: *K-Nearest Neighbor, Prabowo, Sentiment, Term Frequency-Inverse Document Frequency, Twitter.*

I. PENDAHULUAN

Berdasarkan Undang-Undang Dasar 1945 pasal 1 ayat 2, Indonesia merupakan negara demokrasi, dimana kekuasaan tertinggi berasal dari rakyat, oleh rakyat, dan untuk rakyat. Pemilihan presiden (pilpres) melalui pemilihan umum adalah salah satu bentuk demokrasi. Terdapat banyak kandidat calon presiden untuk pilpres 2024 mendatang, salah satu kandidat terkuat adalah Prabowo Subianto. Prabowo Subianto adalah seorang politisi yang sedang menjabat sebagai Menteri Pertahanan Republik Indonesia untuk periode 2019 hingga 2024. Prabowo dipilih sebagai kandidat calon presiden untuk pilpres 2024 mendatang oleh Partai Gerakan Indonesia Raya (Gerindra). Ini merupakan pencalonan ketiga Prabowo sebagai kandidat calon presiden setelah pilpres 2014 dan 2019. Oleh karena itu, Prabowo sebaiknya tahu bagaimana sentimen publik terhadap dirinya guna bisa memperbaiki citra diri menjadi lebih baik lagi. Salah satunya adalah dengan melihat sentimen publik pada media sosial.

Twitter merupakan salah satu media sosial yang paling banyak digunakan dalam mengemukakan sentimen terhadap suatu topik bahasan. Twitter juga banyak digunakan sebagai alat promosi dan kampanye dalam bidang politik. Twitter juga menyediakan fungsi *Application Programming Interface* (API) yang dapat digunakan sebagai alat bertukar data dari suatu aplikasi ke aplikasi lainnya (Santoso, 2021). Hal ini menjadikan Twitter sebagai media yang paling banyak diminati untuk melihat bagaimana sentimen publik terhadap suatu topik, apakah cenderung positif atau negatif. Namun sering terjadi kesulitan dalam mengklasifikasikan sentimen-sentimen tersebut jika dilakukan secara manual karena diperlukan waktu dan usaha yang cukup besar. Oleh karena itu, perlu dilakukan suatu tindakan untuk menentukan kelas-kelas sentimen tersebut dengan menggunakan analisis sentimen.

Analisis sentimen merupakan metode yang digunakan untuk memahami, mengeskstrak, dan mengolah data tekstual seacar otomatis untuk memperoleh sebuah informasi (Arisandi dkk, 2023). Metode yang digunakan untuk mengklasifikasikan sentimen ini adalah metode klasifikasi. Terdapat banyak algoritma dalam metode klasifikasi, salah satunya adalah algoritma *K-Nearest Neighbor.* Kelebihan dari algoritma ini adalah efektif pada data yang memiliki data latih yang besar (Han & Kamber, 2006). Hasil penelitian Zuhdi dkk (2019) terkait sentimen terhadap capres 2019 menggunakan algoritma *K-Nearest Neighbor* (KNN) menghasilkan tingkat akurasi sebesar 83,33%. Penelitian lainnya oleh Prasetyo dkk (2023) terkait sentimen terhadap relokasi Ibu Kota Nusantara menggunakan algoritma

DOI: https://doi.org/10.24036/ujsds/vol1-iss5/101

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Naive Bayes dan KNN menghasilkan bahwa tingkat akurasi dengan algoritma KNN lebih tinggi dibanding algoritma *Naive Bayes* yaitu sebesar 88,12% dengan algoritma KNN dan 82,27% dengan algoritma *Naive Bayes.*

Adapun tujuan dari penelitian ini yaitu untuk mengetahui bagaimana kecenderungan sentimen publik terhadap sosok Prabowo Subianto sebagai capres 2024 mendatang guna bisa dijadikan sebagai referensi agar Prabowo dapat menjadi tokoh publik yang lebih baik lagi selaku capres 2024.

II. METODE PENELITIAN

A. Tahapan Analisis Data

Analisis sentimen digunakan untuk mengelompokkan teks yang ada pada suatu sentimen untuk menentukan apakah sentimen tersebut bersifat positif atau negatif (Liu, 2010). Analisis sentimen dilakukan dengan beberapa tahap, yaitu.

1. Mengumpulkan data dengan meng-*crawling* data Twitter menggunakan fungsi API yang telah disediakan oleh pihak Twitter. Dalam tahapan ini, digunakan kata kunci untuk mendapatkan sentimen terkait topik yang diinginkan untuk dianalisis.
2. *Text preprocessing,* tahapan ini digunakan untuk mengubah data yang tidak testruktur menjadi data yang terstruktur dengan tahapan-tahapan seperti *cleaning data*, meghapus kata/simbol yang tidak penting seperti *retweet, username, hashtag,* dan URL. *Case Folding,* menyamakan ukuran semua huruf pada data. *Removal duplicates,* menghapus data yang memiliki duplikat. *Tokenizing,* memisahkan kalimat menjadi kata-kata. *Stemming,* mendapatkan kata-kata dasar dengan cara menghilangkan awalan dan akhiran. *Stopwords removal,*menghapus kata penghubung dan kata ganti, serta kata keterangan. Dan terakhir, normalisasi kata dengan merujuk pada Kamus Besar Bahasa Indonesia (KBBI) (Zuhdi, dkk, 2019).
3. Pelabelan Data

   Proses pelabelan dilakukan secara otomatis dengan metode *lexicon based.* Kata-kata pada sentimen akan dicocokkan dengan kata-kata yang ada pada kamus *lexicon,* kemudian akan dilakukan perhitungan untuk melihat nilai skor pada sentimen tersebut. Jika skor suatu sentimen besar sama dengan 0, maka sentimen akan dikategorikan ke dalam kelas positif. Dan sebaliknya, jika skor suatu sentimen lebih kecil dari 0, maka sentimen tersebut akan dikategorikan sebagai sentimen negatif (Diwandanu & Wisudawati, 2023). Skor sentimen diperoleh melalui persamaan (1) sebagai berikut.

$$skor = (\sum skor\ positif) - (\sum skor\ negatif) \tag{1}$$

4. Klasifikasi Sentimen dengan Algoritma *K-Nearest Neighbor*

   K-Nearest Neighbor (KNN) adalah algoritma klasifikasi yang berdasarkan pada jarak terdekat suatu data yang akan diuji dengan data latih pada nilai k tertentu (Deviyanto & Wahyudi, 2018). Nilai k menyatakan jumlah tetangga (*neighbor*) terdekat yang digunakan untuk memprediksi kelas dari data yang diuji. Untuk menentukan jarak terdekat tersebut, dilakukanlah pembagian data menjadi data latih dan data uji terlebih dahulu. Adapun langkah-langkah analisis sentimen menggunakan algoritma KNN menurut Deviyanto & Wahyudi (2018) adalah sebagai berikut.

   1. Menghitung bobot setiap kata menggunakan metode TF-IDF.

      Term frequency-inverse document frequency (TF-IDF) merupakan suatu metode pembobotan yang digunakan dalam ekstraksi data teks. *Term frequency* merupakan frekuensi munculnya suatu *term* dalam suatu sentimen. Sedangkan *inverse document frequency* merupakan ukuran pentingnya suatu *term* dalam keseluruhan sentimen yang dihitung sebagai logaritma dari jumlah keseluruhan sentimen dibagi dengan jumlah sentimen yang berisi *term* tersebut (Prasetyo dkk, 2023). Adapun rumus perhitungan untuk TF-Idf dapat dilihat pada persamaan (2) dan (3) sebagai berikut.

$$w_{ij} = tf_{ij} \times idf \tag{2}$$

$$idf = log\left(\frac{N}{df_i}\right) \tag{3}$$

      Dimana w_{ij} merupakan bobot dari kata i pada sentimen ke-j, tf merupakan jumlah kemunculan kata i pada sentimen ke-j, dan df merupakan jumlah sentimen yang mengandung kata i.

   2. Menghitung jarak atau tingkat kemiripan (*Cosine Similarity).*

      Pada tahap ini, akan dihitung jarak atau tingkat kemiripan data dengan setiap data *training* yang ada menggunakan rumus *cosine similarity.* Adapan langkah-langkahnya yaitu.

DOI: https://doi.org/10.24036/ujsds/vol1-iss5/101

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

a) Mengalikan bobot dari setiap kata pada S_i dengan bobot setiap kata dari semua sentimen pada data *training*, kemudian dijumlahkan.
b) Menghitung hasil kuadrat dari bobot masing-masing kata dalam setiap sentimen pada data *training* dan S_i, kemudian jumlahkan lalu diakarkan.
c) Menghitung nilai *cosine similarity* dengan rumus pada persamaan (4).

$$Cos(\theta_{SD}) = \frac{\sum_{i=1}^{n} S_i D_i}{\sqrt{\sum_{i=1}^{n} (S_i)^2} \times \sqrt{\sum_{i=1}^{n} (D_i)^2}} \quad (4)$$

Dimana $Cos(\theta_{SD})$ menyatakan tingkat kemiripan S_i terhadap sentimen ke-i pada data *training*, S_i menyatakan bobot sentimen yang akan diuji, D_i menyatakan bobot sentimen ke-i pada data *training*, dan n menyatakan banyaknya sentimen (Diwandanu & Wisudawati, 2023).

3. Menentukan kelas sentimen dengan mengurutkan nilai *cosine similarity* dari yang tertinggi hingga terendah, kemudian memilih nilai k untuk melihat kelas sentimen yang paling banyak muncul.

5. Evaluasi Model Klasifikasi

Evaluasi model klasifikasi digunakan untuk mengevaluasi suatu model klasifikasi yang diperoleh menggunakan data *testing* yang label kelasnya telah diprediksi oleh model. Evaluasi model ini dilakukan dengan menggunakan suatu *confusion matrix* yang berisikan informasi mengenai kelas aktual dan kelas prediksi yang diperoleh dari data *testing* (Suyanto, 2022). Secara umum, *confusion matrix* memiliki struktur seperti pada Tabel 1 berikut.

Tabel 1. *Confusion Matrix*

Kelas Aktual	**Kelas Prediksi**	
	Positif	Negatif
Positif	*True Positive* (TP)	*False Negative* (FN)
Negatif	*False Positive* (FP)	*True Negative* (TN)

Dimana *True Positive* (TP) menyatakan jumlah sentimen dari kelas positif yang benar diklasifikasikan sebagai kelas positif. *True Negative* (TN) menyatakan jumlah sentimen dari kelas negatif yang benar diklasifikasikan sebagai kelas negatif. *False Positive* (FP) menyatakan jumlah sentimen dari kelas negatif yang salah diklasifikasikan sebagai kelas positif, Dan *False Negative* (FN) menyatakan jumlah sentimen dari kelas positif yang salah diklasifikasikan sebagai kelas negatif.

Adapun persaman yang digunakan untuk menghitung nilai akurasi adalah sebagai berikut.

$$Akurasi = \frac{TP+TN}{TP+FP+FN+T} \times 100\% \quad (5)$$

B. Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh melalui media sosial Twitter dengan kata kunci ”prabowo”. Pengambilan data dimulai pada tanggal 1 April 2023 hingga 30 April 2023 sebanyak 3663 sentimen. Adapun variabel penelitian yang digunakan dapat dilihat pada Tabel 2 sebagai berikut.

Tabel 2. Variabel Penelitian

Variabel	**Keterangan**
Y	Kelas Sentimen (Positif atau Negatif) Skor < 0, sentimen negatif Skor ≥ 0, sentimen positif
X	Frekuensi kata i yang muncul pada sentimen

DOI: https://doi.org/10.24036/ujsds/vol1-iss5/101

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

III. HASIL DAN PEMBAHASAN

A. *Text Preprocessing*

Setelah data diperoleh, selanjutnya dilakukan *text preprocessing* untuk memperoleh data yang terstruktur dengan tahapan *cleaning data*, *case folding, tokenizing, stemming, stopwords removal,* dan normalisasi kata. Berikut hasil proses dari *text preprocessing* yang ditunjukkan pada Tabel 3.

Tabel 3. Contoh Hasil *Text Preprocessimg*

Tweet	Keterangan
@DeenAwwalu @Restty_cayah @prabowo Prabowo mang yg terbaik. Dukukng Parbowo jadi Presiden 2024.	*Tweet* Asli
Prabowo mang yg terbaik Dukukng Parbowo jadi Presiden 2024	*Cleaning Data*
prabowo mang yg terbaik. dukukng parbowo jadi presiden 2024.	*Case Folding*
'prabowo', 'mang', 'yg', 'terbaik', 'dukukng', 'parbowo', 'jadi', 'presiden', '2024'	*Tokenizing*
'prabowo', 'mang', 'yg', 'baik', 'dukukng', 'parbowo', 'jadi', 'presiden', '2024'	*Stemming*
'prabowo', 'baik', 'dukukng', 'parbowo', 'presiden', '2024'	*Stopwords Removal*
'prabowo', 'baik', 'dukung', 'prabowo', 'presiden', '2024'	Normalisasi Kata

Dari total 3663 sentimen yang ada, dilakukanlah tahapan *removal duplicates* terlebih dahulu untuk menghapus data yang memiliki duplikat, sehingga didapatkan data bersih sebanyak 2807 sentimen. Selanjutnya dilakukan proses *text preprocessing* dengan tahapan-tahapan seperti pada Tabel 3 hingga didapatkan kata-kata yang akan dihitung masing-masing bobotnya untuk digunakan pada proses analisis.

B. Pelabelan Data

Setelah didapatkan data yang terstruktur, selanjutnya data akan dilabeli dengan menggunakan metode *lexicon based.* Contoh hasil perhitungan pelabelan data dapat dilihat pada Tabel 4 sebagai berikut.

Tabel 4. Contoh Perhitungan Pelabelan Data

No.	*Tweet*	Kamus Positif		Kamus Negatif		Skor	Kelas Sentimen
1.	prabowo baik dukung prabowo presiden 2024	baik dukung	2	-	0	2	Positif
2.	prabowo gagal capres sekian kali malu capres gagal buat susah pilpres	-	0	gagal malu susah	3	-3	Negatif
3.	senang sangat lihat prabowo sambut megah sumbar dukung pilpres	senang megah dukung	3	-	0	3	Positif
4.	prabowo dukung pilih beliau kalah semoga kali menang aamiin	dukung menang	2	kalah	1	1	Positif

Berikut hasil proses pelabelan data *tweet* secara otomatis oleh metode *lexicon based* yang ditunjukkan pada Tabel 5.

Tabel 5. Pelabelan Data

Kelas	Jumlah
Positif	2731
Negatif	76

DOI: https://doi.org/10.24036/ujsds/vol1-iss5/101

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Berdasarkan Tabel 5, dari 2807 sentimen yang ada, diperoleh sentimen positif sebanyak 2731 sentimen dan sentimen negatif sebanyak 76 sentimen.

C. Klasifikasi dengan Algoritma *K-Nearest Neighbor*

Sebelum melakukan pengklasifikasian dengan algoritma *K-Nearest Neighbor* (KNN), dilakukanlah pembobotan masing-masing kata terlebih dahulu dengan menggunakan metode TF-IDF. Berikut contoh pembobotan kata pada *tweet* nomor 1, 2, 3, dan 4 pada Tabel 4 dengan metode TF-IDF yang ditunjukkan pada Tabel 5.

Tabel 5. Contoh Pembobotan TF-IDF

Kata	tf				df	idf	Bobot (w_i)			
	S_1	D_2	D_3	D_4			w_{S_1}	w_{D_2}	w_{D_3}	w_{D_4}
prabowo	2	1	1	1	4	0	0	0	0	0
baik	1	0	0	0	1	0,602059991	0,602059991	0	0	0
dukung	1	0	1	1	3	0,124938737	0,124938737	0	0,124938737	0,124938737
presiden	1	0	0	0	1	0,602059991	0,602059991	0	0	0
2024	1	0	0	0	1	0,602059991	0,602059991	0	0	0
gagal	0	2	0	0	1	0,602059991	0	1,20411998	0	0
capres	0	2	0	0	1	0,602059991	0	1,20411998	0	0
sekian	0	1	0	0	1	0,602059991	0	0,602059991	0	0
kali	0	1	0	1	2	0,301029996	0	0,602059991	0	0
malu	0	1	0	0	1	0,602059991	0	0,602059991	0	0
buat	0	1	0	0	1	0,602059991	0	0,602059991	0	0
susah	0	1	0	0	1	0,602059991	0	0,602059991	0	0
senang	0	0	1	0	1	0,602059991	0	0	0,602059991	0
sangat	0	0	1	0	1	0,602059991	0	0	0,602059991	0
lihat	0	0	1	0	1	0,602059991	0	0	0,602059991	0
sambut	0	0	1	0	1	0,602059991	0	0	0,602059991	0
megah	0	0	1	0	1	0,602059991	0	0	0,602059991	0
sumbar	0	0	1	0	1	0,602059991	0	0	0,602059991	0
pilpres	0	1	1	0	2	0,301029996	0	0,301029996	0,301029996	0
pilih	0	0	0	1	1	0,602059991	0	0	0	0,602059991
beliau	0	0	0	1	1	0,602059991	0	0	0	0,602059991
kalah	0	0	0	1	1	0,602059991	0	0	0	0,602059991
semoga	0	0	0	1	1	0,602059991	0	0	0	0,602059991
menang	0	0	0	1	1	0,602059991	0	0	0	0,602059991
aamiin	0	0	0	1	1	0,602059991	0	0	0	0,602059991

Setelah diperoleh bobot dari masing-masing kata pada sentimen yang ditunjukkan pada Tabel 5, selanjutnya dilakukan perhitungan *cosine similarity* yang diperoleh menggunakan rumus pada persamaan (4). Sebelumnya, data dipisahkan terlebih dahulu menjadi data *training* dan data *testing.* Pada contoh ini, *tweet* nomor 1 akan menjadi data *testing* yang akan diprediksi label kelasnya, sedangkan *tweet* nomor 2, 3, dm 4 akan menjadi data *training* yang akan digunakan untuk memprediksi label kelas untuk *tweet* nomor 1. Berikut hasil perhitungan *cosine similarity* dari S_1 yang ditunjukkan pada Tabel 6.

Tabel 6. Hasil Perhitungan *Cosine Similarity*

Cosine Similarity		
$Cos(S_1D_2)$	$Cos(S_1D_3)$	$Cos(S_1D_4)$
0	0,009840747	0,023747587

DOI: https://doi.org/10.24036/ujsds/vol1-iss5/101

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Hasil *cosine similarity* dari S_1 yang diperoleh pada Tabel 6 selanjutnya diurutkan dari yang tertinggi ke yang terendah. Didaparkan hasil urutan sebagai berikut.

1. D_4 (Sentimen Positif)
2. D_3 (Sentimen Positif)
3. D_2 (Sentimen Negatif)

Jika nilai k yang dipilih untuk KNN adalah 3, maka dari 3 nilai *cosine similarity* tersebut diperoleh bahwa kelas sentimen yang paling banyak muncul adalah sentimen positif. Dengan demikian, dapat disimpulkan bahwa contoh tersebut masuk ke dalam kategori sentimen positif.

D. Evaluasi Model Klasifikasi

Untuk mengukur ketepatan model dalam mengklasifikasikan sentimen, digunakan nilai akurasi dari sebuah *confusion matrix*. Berikut *confusion matrix* untuk pembagian data *training* dan data *testing* 80:20.

Tabel 7. *Confusion Matrix*

Kelas Aktual	Kelas Prediksi	
	Positif	Negatif
Positif	546	2
Negatif	13	1

Berdasarkan Tabel 7, dapat dilihat bahwa model mengklasifikasikan sentimen dengan benar sebanyak 546 sentimen untuk kelas positif dan 1 sentimen untuk kelas negatif. Dan juga model memprediksi 13 sentimen negatif sebagai sentimen positif dan 2 sentimen positif sebagai sentimen negatif. Adapun nilai akurasi dari model dapat dihitung menggunakan persamaan (5) sebagai berikut.

$$Akurasi = \frac{546 + 1}{546 + 1 + 2 + 13} \times 100\% = 97{,}33\%$$

Dari hasil perhitungan, diiperoleh nilai akurasi dari pengklasifikasian sentimen terkait Prabowo sebagai capres 2024 dengan k = 3 pada pemisahan data *training* dan data *testing* sebesar 80:20 adalah 97,33%.

IV. KESIMPULAN

Berdasarkan hasil pembahasan di atas, diperoleh hasil pengklasifikasian sentimen terhadap Prabowo Subianto sebagai capres 2024 sebanyak 2731 sentimen positif dan 76 sentimen negatif. Ini menunjukkan bahwa kecenderungan sentimen publik terhadap Prabowo Subianto selaku capres 2024 adalah positif. Adapun tingkat akurasi model dalam mengklasifikasikan sentimen menggunakan algoritma *K-Nearest Neighbor* (KNN) didapatkan sebesar 97,33%. Penelitian selanjutnya dapat dilakukan dengan algoritma klasifikasi lainnya seperti *Naive Bayes, Support Vector Machine, Decision Tree, Random Forest,* dan lainnya.

DAFTAR PUSTAKA

Arisandi, R. R., Sumarno, S., dan Setiawan, H. 2023. Comment Sentiment Analysis of JNE Using K-Nearest Neighbor (KNN) Method on Twitter. *Indonesian Journal of Inovation Studies.* 24. 1-21.

Deviyanto, A., dan Wahyudi, M. D. R. 2018. Penerapan Analisis Sentimen pada Pengguna Twitter Menggunakan Metode K-Nearest Neighbor. *Jurnal Informatika Sunan Kalijaga,* 3(1), 1-14.

Diwandanu, M. T. dan Wisudawati, L. M. 2023. Analisis Sentimen Terhadap Twit Maxim pada Twitter Menggunakan R Programming dan K-Nearest Neighbor. *Jurnal Ilmiah Informatika Komputer,* 28(1), 1-16.

Han, J. dan Kamber, M. 2006. *Data Mining: Concepts and Techniques Third.,* Elsevier.

DOI: https://doi.org/10.24036/ujsds/vol1-iss5/101

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Han, J., Kamber, M., dan Pei, J. 2012. *Data Mining Concepts and Techniques Third Edition.* Waltham USA: Morgan Kaufmann.

Liu, B. 2010. *Handbook of Natural Language Processing 2nd Edtion*. Boca Raton: CRC Press.

Santoso, G. T. 2021. *Analisis Sentimen pada Tweet dengan Tagar #bpjsrasarentenir Menggunakan Metode Support Vector Machine (SVM)*. Pekanbaru: Universitas Islam Riau.

Prasetyo, S. D., Hilabi. S. S., dan Nurapriani, F. 2023. Analisis Sentimen Relokasi Ibukota Nusantara Menggunakan Algoritma Naive Bayes dan KNN. *Jurnal KomtekInfo,* 10(1), 1-7.

Suyanto. 2022. *Machine Learning Tingkat Dasar dan Lanjut*. Bandung: INFORMATIKA.

Zuhdi, A. M., Utami, E., dan Raharjo, S. 2019. Analisis Sentiment Twitter Terhadap Capres Indonesia 2019 dengan Metode KNN. *Jurnal INFORMA Politeknik Indonesia Surakarta ISSN,* 5(2).

DOI: https://doi.org/10.24036/ujsds/vol1-iss5/101