

Analisis Sentimen Kepuasan Pengguna Terhadap Layanan Streaming Mola Menggunakan Algoritma *Random Forest*

Setya Nanda¹, Desti Mualfah², Diah Angraina Fitri³

^{1,2,3}Teknik Informatika, Ilmu Komputer, Universitas Muhammadiyah Riau

¹180401056@student.umri.ac.id, ²destimualfah@umri.ac.id, ³diahangrainafitri@umri.ac.id

ABSTRAK

Pada masa sekarang ini perkembangan teknologi berkembang sangat cepat, salah satunya dalam bidang penyediaan informasi, teknologi informasi dapat digunakan untuk melengkapi sebuah data dan biasa digunakan sebagai dasar dalam pengambilan sebuah keputusan. *Platform streaming* yang digandrungi masyarakat saat ini yaitu MOLA. MOLA merupakan aplikasi *platform streaming* video yang menyajikan program hiburan dan olahraga. Aplikasi MOLA dirilis di *Google Play Store* pada tanggal 19 Juli 2019. Namun dengan banyaknya aplikasi serupa yang tersedia tidak menutup kemungkinan para pengguna untuk berkomentar bebas mengeluarkan pendapatnya melalui fitur ulasan yang tersedia di *playstore* dengan banyaknya ulasan komentar peneliti membuat sistem analisis sentimen yang bertujuan untuk mengelompokan ulasan komentar pengguna dalam bentuk sentimen positif negatif dan netral. nertal dengan menggunakan Algoritma *Random Forest* dan menggunakan 520 dataset dengan teknik *scrapping* per tanggal 06 Mei 2022 sampai dengan 22 Juli 2022. Mendapatkan hasil klasifikasi pada kelas positif dengan jumlah 126, negatif 196 dan netral 198. Pembagian data *training* 90% dan *testing* 10% menghasilkan nilai *accuracy* 98%, nilai *precision* 1,00%, nilai *recall* 95% dan nilai *f1-score* 98%.

Kata kunci: *Mola, platform streaming, playstore, sentimen, random forest*

ABSTRACT

At this time the development of technology is growing very fast, one of which is in the field of providing information, information technology can be used to complete a data and is commonly used as a basis for making a decision. The streaming platform that is loved by the community today is Mola. Mola is a video streaming platform application that provides entertainment and sports programs. The Mola application was released on the Google Play Store on July 19, 2019. However, with the number of similar applications available, it is possible for users to freely express their opinions through the review feature available on the playstore. user comments reviews in the form of positive negative and neutral sentiments. neutral using the Random Forest Algorithm and using 520 datasets with scrapping techniques as of 06 May 2022 to 22 July 2022. Obtaining classification results in the positive class with a total of 126, negative 196 and neutral 198. The distribution of 90% training data and 10% testing resulted in the accuracy value is 98%, the precision value is 1.00%, the recall value is 95% and the f1-score value is 98%.

Keywords: *Mola, streaming platform, playstore, sentiment, random forest*

1. PENDAHULUAN

Dalam *Appstore* terdapat informasi-informasi berupa deskripsi, komentar dari user dan rating mengenai aplikasi di dalamnya dengan tujuan agar mengetahui kekurangan ataupun kelebihan dari aplikasi yang dibuat[1]. *Platform streaming* yang digandrungi masyarakat saat ini yaitu MOLA. MOLA merupakan aplikasi *platform streaming* video yang menyajikan program hiburan dan olahraga. Aplikasi MOLA dirilis di *Google Play Store* pada tanggal 19 Juli 2019. Aplikasi mola sudah diunduh sebanyak 5 juta kali namun hanya memiliki rating 3,2. Pada *google play store* pemberian rating aplikasi diikuti dengan ulasan dari para pengguna terhadap aplikasi[2].

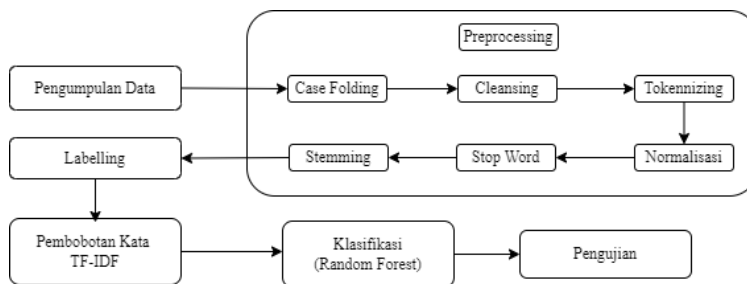
Berdasarkan masalah dan penelitian terdahulu yang telah di paparkan algoritma *Random Forest* akan digunakan untuk melakukan analisis sentimen MOLA dimana data diperoleh dari *playstore*. Penerapan algoritma *Random Forest* pada penelitian digunakan untuk mencari nilai positif negatif dan netral sehingga dapat mengetahui kepuasan pengguna terhadap aplikasi MOLA pada *playstore*. [3]

Penelitian ini menggunakan metode *Random Forest* Pada tahap ini mesin akan diajari untuk mengenal pola atau dokumen yang ada agar dapat mengklasifikasi sebuah data ke dalam dua kelas, yaitu kelas positif dan kelas negatif. Data yang sudah melewati text preprocessing akan dibagi menjadi 2 bagian yaitu data training dan data testing [4]. hasil klasifikasi pada kelas positif dengan jumlah 126, negatif 196 dan netral 198. Pembagian data training 80% dan testing 20% menghasilkan nilai *accuracy* 1,00%, nilai *precision* 1,00%, nilai *recall* 1,00% dan nilai *f1-score* 1,00%.

Hasil dari analisis sentimen ini adalah tentang pendapat dan opini pengguna terhadap aplikasi MOLA dengan cara melakukan klasifikasi kedalam 3 kelas yaitu positif negatif dan netral dengan menggunakan Algoritma *Random Forest* dan menggunakan 520 dataset dengan teknik *scrapping* per tanggal 06 Mei 2022 sampai dengan 22 Juli 2022. Mendapatkan hasil klasifikasi pada kelas positif dengan jumlah 126, negatif 196 dan netral 198.

2. METODOLOGI PENELITIAN

Metodologi dalam sebuah penelitian sangat perlu karena metode penelitian merupakan proses dalam melakukan sebuah penelitian yang merupakan sebuah pedoman dalam mendapatkan hasil yang sesuai dengan tujuan penelitian. Berikut tahapan-tahapan pada penelitian [5].



Gambar 1. Metodologi Penelitian

2.1. Pengumpulan Data

Pada tahap pengumpulan dataset yang digunakan dalam penelitian ini berasal dari *Playstore*, dimana merupakan layanan distribusi digital yang dioperasikan sekaligus dikembangkan oleh *Google*. *Playstore* berfungsi sebagai toko aplikasi resmi buat seluruh perangkat yang bersistem operasi Android. *Google Play* ini bisa menelusuri dan mengunduh aplikasi yang dikembangkan dengan Android SDK atau *Software Development Kit* yang sudah diterbitkan oleh *Google*. Pada pengambilan data dari *Playstore* menggunakan metode *Webscrapping* dengan tools *Google Colab* dengan bahasa pemrograman *Python* [6]. Dengan cara *scraping* data terbaru.

2.2. Preprocessing

Preprocessing merupakan tahapan awal yang dapat dilakukan untuk mempermudah dalam pengambilan maupun menerapkan pengambilan data dimana data yang kurang sempurna maupun data yang tidak konsisten dapat disaring menjadi data yang diperlukan. Berikut terdapat beberapa tahapan-tahapan *preprocessing* data :

2.1.1 Case Folding

Tahapan *case folding* mengkonversikan keseluruhan teks dari huruf capital menjadi huruf kecil. Hanya huruf “a” sampai “z” yang akan diterima dalam proses ini selain huruf tersebut semuanya dianggap tidak diterima.

2.1.2 Cleaning

Tahapan *cleaning* merupakan suatu proses pembersihan dokumen dari kata atau karakter yang dihilangkan pada komentar adalah karakter simbol, angka, link, url, hastag (#), mention (@username), tanda koma, titik dan tanda baca lainnya. *Tokenizing*

2.1.3 Tokenizing

Tahapan *tokenization/tokenizing* merupakan pemilahan kalimat menjadi perkata. Suatu proses pemotongan seluruh karakter dan sebuah set dokumen yang diberikan menjadi satu potongan kata. Potongan tersebut dikenal dengan istilah token.

2.1.4 Normalisasi

Tahapan *Normalisasi* adalah teknik yang sering diterapkan sebagai bagian dari penyiapan data untuk pembelajaran mesin. *Normalisasi* juga diperlukan untuk beberapa algoritma untuk pembuatan model.

2.1.5 Stop Removal

Tahapan *stop removal* merupakan kata umum yang biasanya muncul dalam jumlah besar dan dianggap tidak memiliki makna. Seperti kata konjungsi “yang”, “dan”, “dari”, dll. Atau sebagai penghapus kata penghubung dan imbuhan.

2.1.6 Stemming

Tahapan *stemming* merupakan proses membentuk suatu kata menjadi kata dasar. Proses ini menghilangkan semua imbuhan.

2.3. Pembobotan Kata TF-IDF

Data yang sudah melalui proses text preprocessing selanjutnya dihitung berapa banyak frekuensi kemunculan setiap kata didalam dokumen[7]. Nilai TF-IDF dari sebuah kata merupakan kombinasi dari nilai *tf* dan nilai *idf* dalam perhitungan bobot Metode[8].

Tabel 1. Hasil Pelabelan

No	Content	Hasil	Sentimen
0	mantab gambar bagus suara juga dolby atmos 	[mantap, gambar, bagus, suara, juga, dolby, atmos]	Positif
1	Tingkatkan lagi	[tingkatkan]	Netral
2	Coba dl	[coba]	Netral
3	Sy baru coba membukanya	[saya, baru, coba, membukanya]	Netral

2.4. Pengujian

Pada tahap pengujian dilakukan untuk menguji kinerja dari mesin yang telah dibangun, untuk menghitung akurasi dan mengidentifikasi klasifikasi algoritma random forest menggunakan *confusion matrix* dengan menghitung akurasi, *precision*, *recall* dan *F1-Score*.

3. HASIL DAN PEMBAHASAN

Berdasarkan yang telah dijelaskan pada bab sebelumnya, data didapatkan dari ulasan pengguna aplikasi Mola di Play Store. Pengumpulan data dilakukan menggunakan *google colab* dengan memanfaatkan *tools python* dan menginstall beberapa library yang dibutuhkan.

3.1 Pengumpulan Data

Pengumpulan data dilakukan dengan proses scrapping dengan jumlah data 5.000 ulasan pengguna aplikasi Mola. Data yang di scrapping ini nantinya akan di dowload dalam bentuk Microsoft Exel. Sebelum melakukan proses scrapping proses pertama yaitu menginstal goolge play scraper pada Tools Python seperti gambar berikut.

MENGINSTAL GOOGLE PLAY SCRAPER

```
[ ] !pip install google-play-scraper

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting google-play-scraper
  Downloading google_play_scraper-1.2.0-py3-none-any.whl (28 kB)
Installing collected packages: google-play-scraper
Successfully installed google-play-scraper-1.2.0

[ ] from google_play_scraper import app

import pandas as pd

import numpy as np
```

Gambar 2. Perintah Instal *Google Scrap*

3.1.1 Pelabelan

Setelah dataset berhasil di *scrapping* dapat dilihat jumlah teratas dan terbawahnya kemudian dilakukan tahapan berikutnya yaitu pelabelan positif dan negatif . Dalam proses ini data yang sebelumnya masih bentuk mentahan kemudian di label menjadi dua bagian yaitu *Positive* dan *Negative*. Berikut data ditunjukkan pada gambar dibawah.

	userName	score	at	content	klasifikasi
0	Movie Collection	5	2022-07-20 14:01:45	mantab gambar bagus suara juga dolby atmos 👍	Positive
1	ari Kusmawan	2	2022-07-19 13:16:33	Tingkatkan lagi	Negative
2	Y B	4	2022-07-19 10:24:58	Coba di	Positive
3	Alibaba Nabi	5	2022-07-19 10:23:40	Sy baru coba membukanya	Positive
4	Orang Biasa	1	2022-07-17 14:39:01	FREE EPL	Negative

Gambar 3. Hasil Proses Pelabelan

3.1.2 Preprocessing



Tahapan *preprocessing* dilakukan untuk membersihkan data dari hasil *scrapping*. Data *preprocessing* merupakan salah satu tahapan dalam melakukan mining data sebelum menuju ke

tahap pemrosesan, data mentah akan diolah terlebih dahulu. Karena data hasil *scrapping* masih tidak terstruktur diubah menjadi data yang dapat dianalisis atau mudah diolah oleh sistem.

3.1.3 Case Folding

Hasil dari tahapan *case folding*. Menggunakan *library lower* mengubah huruf capital yang ada didokumen menjadi huruf kecil. Berikut hasil *case folding* terdapat pada tabel dibawah .


Tabel 2. Hasil Case Folding

No	Content	Hasil
0	mantap gambar bagus suara juga dolby atmos 	mantap gambar bagus suara juga dolby atmos 
1	Tingkatkan lagi	tingkatkan lagi
2	Coba dl	coba dl
3	Sy baru coba membukanya	sy baru coba membukanya
4	FREE EPL	free epl
5	Sebelumnya bisa, sekarang gabisa bayar pake pulsa, pulsa tidak mencukupi terus, padahal pulsa melimpah	sebelumnya bisa, sekarang gabisa bayar pake pulsa, pulsa tidak mencukupi terus, padahal pulsa melimpah

3.1.4 Cleansing

Pada tahapan *cleansing* ini adalah untuk menghapus karakter yang tidak memberikan pengaruh terhadap proses klasifikasi sentimen seperti menghapus tanda baca koma (,), titik (.), hastag (#), mention (@), link dan angka. Berikut adalah perintah untuk melakukan tahapan *cleansing*.


Tabel 3. Hasil Cleansing

No	Content	Hasil
0	mantap gambar bagus suara juga dolby atmos 	mantap gambar bagus suara juga dolby atmos
1	Tingkatkan lagi	tingkatkan lagi
2	Coba dl	coba dl
3	Sy baru coba membukanya	sy baru coba membukanya
4	FREE EPL	free epl
5	Sebelumnya bisa, sekarang gabisa bayar pake pulsa, pulsa tidak mencukupi terus, padahal pulsa melimpah	sebelumnya bisa sekarang gabisa bayar pake pulsa pulsa tidak mencukupi terus padahal pulsa melimpah

3.1.5 Tahapan Tokenizing

Tokenizing sendiri adalah metode untuk melakukan pemisahan kata dalam suatu kalimat dengan tujuan untuk proses analisis teks lebih lanjut.


Tabel 4. Hasil Tokenizing

No	Content	Hasil
0	mantab gambar bagus suara juga dolby atmos 	[mantap, gambar, bagus, suara, juga, dolby, atmos]
1	Tingkatkan lagi	[tingkatkan, lagi]
2	Coba dl	[coba, dl]
3	Sy baru coba membukanya	[sy, baru, coba, membukanya]
4	FREE EPL	[free, epl]
5	Sebelumnya bisa, sekarang gabisa bayar pake pulsa, pulsa tidak mencukupi terus, padahal pulsa melimpah	[sebelumnya, bisa, sekarang, gabisa, bayar, pake, pulsa, pulsa, tidak, mencukupi, terus, padahal, pulsa, melimpah]

3.1.6 Normalisasi

Perintah untuk melakukan tahapan normalisasi. Pada tabel dibawah ini merupakan hasil dari tahapan normalisasi.


Tabel 5. Hasil Normalisasi

No	Content	Hasil
0	mantab gambar bagus suara juga dolby atmos 	[mantap, gambar, bagus, suara, juga, dolby, atmos]
1	Tingkatkan lagi	[tingkatkan, lagi]
2	Coba dl	[coba, dulu]
3	Sy baru coba membukanya	[saya, baru, coba, membukanya]
4	FREE EPL	[free, epl]
5	Sebelumnya bisa, sekarang gabisa bayar pake pulsa, pulsa tidak mencukupi terus, padahal pulsa melimpah	[sebelumnya, bisa, sekarang, tidakbisa, bayar, pakai, pulsa, pulsa, tidak, mencukupi, terus, padahal, pulsa, melimpah]

3.1.7 Stopword

Berikut merupakan hasil dari tahapan *stopword* yaitu membersihkan kata penghubung contohnya saya, lagi, dan, aku, yang dan imbuhan serta kata sambung lainnya.

Tabel 6. Hasil Stopword

No	Content	Hasil
0	mantab gambar bagus suara juga dolby atmos 	[mantap, gambar, bagus, suara, juga, dolby, atmos]
1	Tingkatkan lagi	[tingkatkan]
2	Coba dl	[coba]
3	Sy baru coba membukanya	[saya, baru, coba, membukanya]
4	FREE EPL	[free, epl]
5	Sebelumnya bisa, sekarang gabisa bayar pake pulsa, pulsa tidak mencukupi terus, padahal pulsa melimpah	[tidakbisa, bayar, pakai, pulsa, pulsa, tidak, mencukupi, terus, padahal, pulsa, melimpah]

3.1.8 Stemming

Tahapan *Stemming* adalah suatu proses pengembalian suatu kata berimbuhan ke dalam bentuk dasarnya. Proses ini akan menghilangkan awalan, akhiran, sisipan dan confixes (kombinasi

dari awalan dan akhiran). Stemming yang digunakan adalah stemming Sastrawi, yang merupakan stemmer pengembangan dari Algoritma Nazief dan Adriani yang awalnya dibangun pada Bahasa pemrograman PHP, kemudian dikembangkan pada Bahasa pemrograman python, java, C, Go, Ruby. Sastrawi sangat bergantung pada kamus kata dasar yang diambil dari kateglo.com dengan perubahan[9]

3.1.9 Hasil *Preprocessing*

Hasil dari tahapan preprocessing kemudian disimpan dengan format csv. Untuk menampilkan hasil dari tahapan preprocessing dengan perintah seperti pada gambar dibawah ini adalah contoh pengambilan data Preprocessing.[10]

```
dataset['CleanContent'] = dataset['stemming'].apply(lambda x: ' '.join(x))  
dataset.to_csv("datasetclean.csv")  
dataset.head(10)
```

Gambar 4. Perintah Pengambilan Data *Preprocessing*

	userName	score	at	content	klasifikasi	cleansing	tokenize	normalisasi	stopword	stemming	CleanContent
0	Movie Collection	5	2022-07-20 14:01:45	mantab gambar bagus suara juga dolby atmos 🍿	Positive	mantab gambar bagus suara juga dolby atmos	[mantab, gambar, bagus, suara, juga, dolby, at...	[mantab, gambar, bagus, suara, juga, dolby, at...	[mantab, gambar, bagus, suara, dolby, atmos]	[mantab, gambar, bagus, suara, dolby, atmos]	mantab gambar bagus suara dolby atmos
1	ari Kusmawan	2	2022-07-19 13:16:33	tingkatkan lagi	Negative	tingkatkan lagi	[tingkatkan, lagi]	[tingkatkan, lagi]	[tingkatkan]	[tingkat]	tingkat
2	Y B	4	2022-07-19 10:24:58	coba di	Positive	coba di	[coba, di]	[coba, dulu]	[coba]	[coba]	coba
3	Alibaba Nabi	5	2022-07-19 10:23:40	sy baru coba membukanya	Positive	sy baru coba membukanya	[sy, baru, coba, membukanya]	[saya, baru, coba, membukanya]	[coba, membukanya]	[coba, buka]	coba buka
4	Orang Biasa	1	2022-07-17 14:39:01	free epl	Negative	free epl	[free, epl]	[free, epl]	[free, epl]	[free, epl]	free epl
5	Reticulation	1	2022-07-17 00:09:41	sebelumnya bisa, sekarang gabisa bayar pake pu...	Negative	sebelumnya bisa sekarang gabisa bayar pake pu...	[sebelumnya, bisa, sekarang, gabisa, bayar, pa...	[sebelumnya, bisa, sekarang, tidakbisa, bayar, pa...	[tidakbisa, bayar, paka, pulsa, pulsa, mencuk...	[tidakbisa, bayar, paka, pulsa, pulsa, cukup...	tidakbisa bayar paka pulsa pulsa cukup pulsa ...

Gambar 5. Hasil *Preprocessing*

3.2 Pembobotan TF-IDF

Data yang sudah melewati tahap preprocessing tersebut harus berbentuk numerik agar bisa masuk kedalam proses klasifikasi. Data tersebut bisa diubah bentuknya menjadi numerik menggunakan metode pembobotan TF-IDF.[11] Nilai TF-IDF dari sebuah kata merupakan kombinasi dari nilai *tf* dan nilai *idf* dalam perhitungan bobot Metode TF-IDF ini menggabungkan dua konsep yaitu frekuensi kemunculan sebuah kata di dalam sebuah dokumen dan inverse frekuensi dokumen yang mengandung kata tersebut.

```
from sklearn.feature_extraction.text import TfidfVectorizer  
X = dataset['CleanContent'].values.astype('U')  
label = dataset['sentimen']  
tfidf_vectorizer = TfidfVectorizer()  
tfidf_vector = tfidf_vectorizer.fit_transform(X)  
print("Hasil dari Tahap TF-IDF :")  
print(tfidf_vector.shape)
```

Gambar 6. Perintah TF-ID

Pada gambar dibawah ini adalah menampilkan jumlah kata yang paling sering muncul berdasarkan hasil TF-IDF, kata yang berukuran besar adalah kata yang paling sering muncul.

selanjutnya melakukan klasifikasi sentimen menggunakan *Random Forest*[13]. Berikut perintah tahapan *Random Forest* terlihat pada gambar dibawah ini.

```
from sklearn.ensemble import RandomForestClassifier
rfc = RandomForestClassifier(n_estimators=100, random_state=0)
rfc.fit(X_test, y_test)
y_pred_rfc = rfc.predict(X_test) # predicting test data
print("Hasil Dari RFC Score Data Training : ",rfc.score(X_train, y_train))
print("Hasil Dari RFC Score Data Testing : ",rfc.score(X_test, y_test))
```

Gambar 10. Perintah *Random Forest*

3.4 Hasil Pengujian

Pada tahapan pengujian menggunakan confusion matrix untuk mengetahui informasi tentang nilai-nilai yang diprediksi dan hasil yang sebenarnya, biasa digunakan untuk perhitungan accuracy, recall, precision dan F1-score. Confusion matrix menyatakan jumlah data uji yang benar diklasifikasikan dan jumlah data uji yang salah diklasifikasikan. Perintah pengujian Confusion Matrix sebagai berikut.

Tabel 7. Hasil Pengujian

Pembagian	Precision	recall	F1-Score	Suport
Negatif	1.00	0.95	0.98	22
Netral	0.94	1.00	0.97	15
Positif	1.00	1.00	1.00	15
Accuracy			0.98	52
Macro Avg	0.98	0.98	0.98	52
Weighted Avg	0.98	0.98	0.98	52

4. KESIMPULAN

Hasil dari analisis sentimen ini adalah tentang pendapat dan opini pengguna terhadap aplikasi MOLA dengan cara melakukan klasifikasi kedalam 3 kelas yaitu positif negatif dan netral dengan menggunakan Algoritma *Random Forest* dan menggunakan 520 dataset dengan teknik *scrapping* per tanggal 06 Mei 2022 sampai dengan 22 Juli 2022. Mendapatkan hasil klasifikasi pada kelas positif dengan jumlah 126, negatif 196 dan netral 198. Pembagian data training 80% dan testing 20% menghasilkan nilai *accuracy* 98%, nilai *precision* 1,00%, nilai *recall* 95% dan nilai *f1-score* 98%.

DAFTAR PUSTAKA

- [1] G. A. Sandag, U. Klabat, and J. A. Mononutu, "Prediksi Rating Aplikasi App Store Menggunakan Algoritma Random Forest Application Rating Prediction on App Store using Random Forest Algorithm," vol. 6, no. 2, pp. 167–178, 2020.
- [2] U. S. Karawang, "3 1,2,3," vol. 5, 2022.
- [3] R. Amri, "ANALISIS EKONOMI POLITIK KOMUNIKASI PADA HAK SIAR PERTANDINGAN SEPAKBOLA Studi Deskriptif Praktik Spasialisasi Dalam Siaran Langsung Sepakbola Liga Eropa di Indonesia," vol. 2, no. 11, pp. 1–12, 2021.
- [4] A. Deolika, K. Kusriani, and E. T. Luthfi, "Analisis Pembobotan Kata Pada Klasifikasi Text Mining," J. Teknol. Inf., vol. 3, no. 2, p. 179, 2019, doi: 10.36294/jurti.v3i2.1077.

- [5] M. Muhsi, "Model dan Analisa Faktor Eksternal Aktifitas Siswa Kelas X TKJ SMKN 1 Pakong Pamekasan Menggunakan Algoritma Decision Tree," *J. Apl. Teknol. Inf. dan Manaj.*, vol. 2, no. 2, pp. 92–106, 2021, doi: 10.31102/jatim.v2i2.1239.
- [6] V. C. Muller and N. B. A, "Survei Pendapat Ahli," pp. 553–571, 2016.
- [7] I. Afdhal, R. Kurniawan, I. Iskandar, R. Salambue, E. Budianita, and F. Syafrina, "Penerapan Algoritma Random Forest Untuk Analisis Sentimen Komentar Di YouTube Tentang Islamofobia," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 1, pp. 49–54, 2022, [Online]. Available: <http://ojs.serambimekkah.ac.id/jnknti/article/view/4004/pdf>.
- [8] D. Mualfah, W. Fadila, and R. Firdaus, "Jurnal Computer Science and Information Technology (CoSciTech) Teknik SMOTE Untuk Mengatasi Imbalance Data Pada Deteksi Penyakit Stroke Menggunakan Algorithm," vol. x, no. x, pp. 1–8, 2022.
- [9] I. K. Dharmendra, N. Nym, U. Januhari, I. P. Ramayasa, and I. M. A. Wirahadi, "Uji Komparasi Sentiment Analysis Pada Opini Alumni Terhadap Perguruan Tinggi," vol. 07, pp. 1–6, 2022.
- [10] A. K. Fauziyyah, "Analisis Sentimen Pandemi Covid19 Pada Streaming Twitter Dengan Text Mining Python," *J. Ilm. SINUS*, vol. 18, no. 2, p. 31, 2020, doi: 10.30646/sinus.v18i2.491.
- [11] N. Anggraini, E. S. N. Harahap, and T. B. Kurniawan, "Text Mining - Analisis Teks Terkait Isu Vaksinasi COVID-19," *J. Ilmu Pengetah. dan Teknol. Komun.*, vol. 23, no. 2, pp. 141–153, 2021.
- [12] A. Primajaya and B. N. Sari, "Random Forest Algorithm for Prediction of Precipitation," *Indones. J. Artif. Intell. Data Min.*, vol. 1, no. 1, p. 27, 2018, doi: 10.24014/ijaidm.v1i1.4903.
- [13] M. B. Sulthan, I. Wahyudi, and L. Suhartini, "Analisis Sentimen Pada Bencana Alam Menggunakan Deep Neural Network dan Information Gain," *J. Apl. Teknol. Inf. dan Manaj.*, vol. 2, no. 2, pp. 65–71, 2021, doi: 10.31102/jatim.v2i2.1273.