

Analisis Opini Terhadap Aplikasi Riliv di Twitter Menggunakan Algoritma *Naïve Bayes* dan *Random Forest*

*Diana Nurfitriana¹, Taufik Ridwan², Apriade Voutama³

^{1,2,3}Sistem Informasi, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang
Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Kota Karawang, Jawa Barat
Email: ¹diana.nurfitriana19011@student.unsika.ac.id, ²taufik.ridwan@cs.unsika.ac.id,
³apriade.voutama@staff.unsika.ac.id

ABSTRACT

In the current era of technological advancement and the internet, people can easily access various information. This technological advancement brings innovation in the mental health field, such as services in the form of apps. This research conducts sentiment analysis using the Naïve Bayes and Random Forest algorithms. The study aims to analyze Twitter users' opinions regarding the Riliv apps and compare the results of classification using Naïve Bayes and Random Forest. This research methodology uses the AI Project Cycle method. The data used is tweet data from Twitter with the keyword 'aplikasi riliv'. The dataset consisted of 1035 data, which was processed to produce 273 positive, 273 neutral, and 39 negative sentiments data. The Naïve Bayes and Random Forest algorithms were applied to compare the classification results of the two. The most optimal classification results are Naïve Bayes with SMOTE with the division of 90% training data and 10% testing data, which results in an accuracy value of 82.72%, a value of precision is 82.89% and a value of recall is 82.72%. Based on the results of the distribution of sentiment data, most users gave positive reviews and were knowledgeable about the Riliv application, while only a few were disappointed.

Keywords : AI project cycle; naïve bayes; random forest; riliv application; sentiment analysis

ABSTRAK

Di era kemajuan teknologi dan internet saat ini, masyarakat dapat dengan mudah mengakses berbagai informasi. Kemajuan teknologi telah menghadirkan inovasi dalam bidang kesehatan mental, seperti pelayanan dalam bentuk aplikasi. Penelitian ini melakukan analisis sentimen menggunakan algoritma *Naïve Bayes* dan *Random Forest*. Tujuan penelitian ini adalah untuk menganalisis opini pengguna Twitter mengenai aplikasi Riliv dan membandingkan hasil dari klasifikasi menggunakan *Naïve Bayes* dan *Random Forest*. Metodologi penelitian ini menggunakan metode *AI Project Cycle*. Data yang digunakan adalah data *tweet* dari Twitter dengan kata kunci 'aplikasi riliv'. Dataset berjumlah 1035 data diproses menghasilkan data sentimen sebanyak 273 positif, 273 netral, dan 39 negatif. Penerapan algoritma *Naïve Bayes* dan *Random Forest* dilakukan perbandingan hasil klasifikasi dari keduanya. Hasil klasifikasi paling optimal yaitu *Naïve Bayes* dengan SMOTE pada pembagian *data training* 90% dan *data testing* 10%, menghasilkan akurasi sebesar 82.72%, *precision* sebesar 82.89%, *recall* sebesar 82.72%. Berdasarkan hasil persebaran data sentimen, sebagian besar pengguna memberikan ulasan positif dan mengetahui aplikasi Riliv, sedangkan sedikit yang merasa kecewa terhadap aplikasi tersebut.

Kata kunci : AI project cycle; analisis sentimen; aplikasi riliv; naïve bayes; random forest

1. PENDAHULUAN

Pada era modern dengan teknologi yang kian berkembang, teknologi dituntut memberikan kemudahan dalam memenuhi banyak kebutuhan di berbagai aspek kehidupan serta akses informasi yang sebagian besar dapat terpenuhi dengan adanya internet dan teknologi. Dengan berkembangnya teknologi, konsultasi di bidang kesehatan dapat dilakukan secara daring tanpa harus mengunjungi rumah sakit ataupun tempat sejenis (Wardani & Erfina, 2021). Aplikasi Riliv merupakan aplikasi pelopor kesehatan mental di Indonesia yang terlahir dari banyaknya stigma negatif di lingkungan masyarakat berdasarkan observasi oleh CEO Riliv, Audrey Maximillian Herli (Tiara, 2021).

Pada tahun 2019, aplikasi Riliv menerima penghargaan *Google Play Best Unique App* 2019 oleh Google yang merupakan perusahaan raksasa teknologi global. Tim *editor* Google Play memberikan penghargaan tersebut berdasarkan *rating* dari pengguna aplikasi. *Co-founder* sekaligus CEO Riliv mengungkapkan bahwa sampai tahun 2019, lebih dari 150 ribu pengguna Riliv menjadi lebih sehat secara mental dengan bantuan aplikasi Riliv (Wijoto, 2019). Hampir 50% dari pengguna

internet melihat rekomendasi dari pengguna lain yang telah memakai produk atau jasa tertentu sebagai pertimbangan apakah akan menggunakan produknya atau tidak (Saputra et al., 2019). Maka dari itu, opini dari pengguna aplikasi menjadi faktor yang penting untuk meningkatkan kualitas aplikasi.

Media sosial menjadi tempat untuk mengungkapkan opini pengguna terhadap suatu topik karena kebebasan dan kesederhanaannya (Giovani et al., 2020). Twitter merupakan media sosial berbasis mikroblog dengan 280 karakter maksimal dalam satu unggahan *tweet* (Aldean et al., 2022). Dalam menganalisis opini yang terdapat pada *tweet* dapat dilakukan analisis sentimen. Analisis sentimen merupakan metode analisis komentar mengenai suatu isu untuk mengetahui sentimen yang terkandung di dalamnya (Wardani & Erfina, 2021). Penelitian ini melakukan analisis sentimen menggunakan algoritma klasifikasi *Naïve Bayes* dan *Random Forest*. Adanya kemungkinan data yang tidak seimbang akan ditangani dengan metode SMOTE. Hasil analisis sentimen diharapkan dapat bermanfaat bagi calon pengguna yang tertarik menggunakan aplikasi Riliv. Penelitian

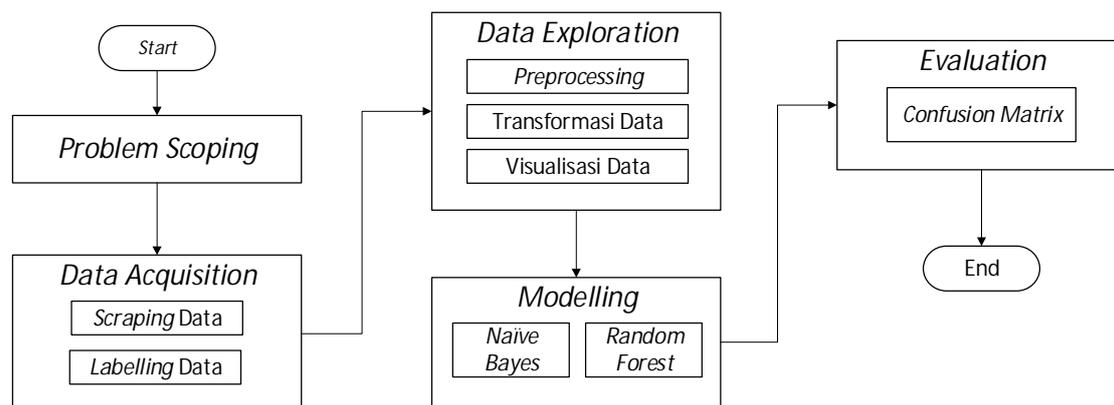
ini juga untuk mengetahui perbandingan performa dari model *Naïve Bayes* dan *Random Forest*.

Penelitian terdahulu tentang analisis sentimen menggunakan algoritma yang diterapkan pada penelitian ini diantaranya, penelitian oleh (Nanda et al., 2022) menggunakan *Random Forest* diperoleh klasifikasi pada kelas positif, negatif dan netral. Hasil evaluasi klasifikasi dengan pembagian data *training* sebesar 90% dan *testing* sebesar 10% menunjukkan nilai akurasi sebesar 98%, presisi 100%, *recall* 95% dan *f1-score* 98%. Penelitian lainnya oleh (Hendra & Fitriyani, 2021) menggunakan *Naïve Bayes*, hasilnya data diklasifikasikan menjadi kelas

positif dan negatif menghasilkan nilai akurasi sebesar 81.68%.

2. METODE

Penelitian ini menggunakan metode *AI Project Cycle*. *AI Project Cycle* merupakan suatu tahapan menggunakan metode ilmiah untuk memecahkan suatu masalah dan menarik kesimpulan mengenai permasalahan (Halim et al., 2022). *AI Project Cycle* memberikan arahan untuk memenuhi tujuan dalam proyek AI menggunakan *framework* atau kerangka kerja yang sesuai. Gambar 1 menunjukkan metode ini memiliki lima tahapan yaitu, *Problem Scoping*, *Data Acquisition*, *Data Exploration*, *Modelling* dan *Evaluation*.



Gambar 1. Rancangan Penelitian

2.1 Tahap *Problem Scoping*

Tahap ini dilakukan identifikasi lingkup permasalahan untuk mempermudah dalam menyelesaikan proyek menggunakan konsep 4W yang

terdiri dari *who*, *what*, *where* dan *why* yaitu, siapa yang berpengaruh terhadap masalah, apa masalah yang terjadi, dimana masalah muncul dan manfaat apa yang akan didapat orang-orang yang berkaitan dari solusi.

2.2 Tahap Data Acquisition

Data penelitian dikumpulkan dari *tweet* di Twitter yang berkaitan dengan aplikasi Riliv. Data dikumpulkan menggunakan *library* python. Data diberi label secara manual, kemudian seleksi data sesuai kebutuhan dari penelitian.

2.3 Tahap Data Exploration

Tahapan ini dilakukan untuk memahami karakteristik dari data dan pola dari data tersebut. Data yang telah dikumpulkan, diatur dan disusun supaya menjadi data yang dapat diolah dengan baik. Tahap ini terdiri dari *preprocessing* data, transformasi data dan visualisasi data.

a. Preprocessing

Preprocessing dilakukan untuk membersihkan dan memproses data teks menjadi data yang lebih terstruktur dari kesalahan dan gangguan (Birjali et al., 2021). Tahapan ini terdiri dari *cleaning*, *case folding*, *tokenizing*, *remove stopword*, *stemming* dan *filter tokens by length*. *Cleaning* merupakan proses pembersihan data dari *noise* seperti tanda baca, *retweet*, *hashtag*, *url*, dan lain-lain. *Case folding* merupakan proses perubahan kata menjadi huruf kecil. *Tokenizing* merupakan proses pemecahan kalimat menjadi per kata.

Remove stopword dilakukan penyaringan kata-kata dalam dokumen yang disebut *stopword*. *Stemming* merupakan proses perubahan kata berimbuhan menjadi kata dasarnya. *Filter tokens by length* dilakukan penghapusan karakter yang berjumlah diluar dari *range* yang ditentukan (Giovani et al., 2020).

b. TF-IDF (*Term Frequency-Inverse Document Frequency*)

TF-IDF adalah proses untuk melakukan pembobotan kata dengan merubah data dari data teks menjadi data numerik. TF-IDF digunakan untuk mengevaluasi seberapa berarti suatu kata yang ada di dokumen. Rumus TF-IDF sebagai berikut (Fadhillah, 2022).

$$tf - idf = tf \times idf \quad (1)$$

$$idf = \log \frac{N}{df} \quad (2)$$

TF atau *Term Frequency* adalah frekuensi kemunculan kata dalam dokumen, IDF atau *Inverse Document Frequency* adalah nilai *inverse* dari DF, DF adalah jumlah dokumen yang mengandung *term* sebuah kata, N adalah jumlah total dokumen dalam corpus N.

2.4 Tahap Modelling

Data diproses menerapkan algoritma *Naïve Bayes* dan *Random Forest* secara terpisah. Data akan dibagi

menjadi data *training* dan data *testing*. Adapun data yang tidak seimbang akan ditangani dengan teknik SMOTE, sehingga data hasil SMOTE tersebut juga akan diproses dengan kedua algoritma.

a. *Naïve Bayes*

Naïve Bayes adalah *classifier* sederhana berdasarkan pada Teorema *Bayes* dan paling umum digunakan dalam klasifikasi teks (Birjali et al., 2021). Konsep yang mendasari aturan *Bayes* yaitu hasil dari hipotesis atas peristiwa (H) yang dapat diperkirakan berdasarkan bukti (E) yang diamati.

$$P(H|E) = \frac{P(E|H) \times P(H)}{P(E)} \quad (3)$$

$P(H|E)$ adalah probabilitas hipotesis H jika diberikan bukti E terjadi, $P(E|H)$ adalah probabilitas bukti E terjadi akan mempengaruhi hipotesis H (*likelihood*), $P(H)$ adalah probabilitas hipotesis H (*prior probability*), dan $P(E)$ adalah probabilitas bukti E (*predictor prior probability*).

b. *Random Forest*

Random Forest merupakan pengembangan dari *Decision Tree* (Aldean et al., 2022). *Random Forest* bersifat *ensemble* yang berarti menggunakan sejumlah *decision tree*

dan menciptakan *forest* yang dapat memperoleh nilai akurasi lebih optimal.

c. SMOTE (*Synthetic Minority Over Sampling Technique*)

SMOTE merupakan teknik *oversampling* yang dilakukan dengan membuat sampel sintetis untuk kelas minoritas, sehingga dapat mengatasi masalah ketidakseimbangan kelas (Aldean et al., 2022). Teknik SMOTE secara acak memilih data dari kelas minoritas yang diidentifikasi ke data terdekat dari data tersebut dan salah satunya dipilih secara acak kemudian dihubungkan dengan data awal untuk membentuk segmen garis di ruang fitur sehingga dihasilkan data sintetis. (Qadrini et al., 2022).

2.5 Tahap *Evaluation*

Model yang telah dibentuk pada tahap *Modelling* dievaluasi menggunakan *confusion matrix* untuk mengetahui performa dari model yang dibuat dengan nilai akurasi, *precision* dan *recall*.

a. *Confusion Matrix*

Confusion matrix adalah tabel untuk mengevaluasi keefektifan model klasifikasi terhadap data *training*. *Confusion matrix* dapat digunakan untuk mengevaluasi model dengan metrik evaluasi akurasi, presisi dan *recall*.

Akurasi adalah metrik untuk menentukan seberapa akurat model. Presisi adalah metrik untuk menunjukkan kekuatan dari prediksi. *Recall* adalah metrik untuk mengukur kesalahan klasifikasi yang dilakukan model (Wankhade et al., 2022).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (6)$$

TP (*True Positif*) adalah jumlah sampel positif yang diprediksi dengan benar sebagai positif, FP (*False Positif*) jumlah sampel negatif yang salah diprediksi sebagai positif, FN (*False Negatif*) adalah jumlah sampel positif yang salah diprediksi sebagai negatif, TN (*True Negatif*) adalah jumlah sampel negatif yang diprediksi dengan benar sebagai negatif.

3. HASIL DAN PEMBAHASAN

Hasil penelitian yang dilakukan adalah mengetahui analisis sentimen dari para pengguna Twitter di Indonesia terhadap aplikasi Riliv dengan menggunakan *Naïve Bayes* dan *Random Forest*, serta mengetahui hasil performa dari model yang telah dibuat.

3.1 Problem Scoping

Pendefinisian lingkup permasalahan dari penelitian yakni mengklasifikasikan opini yang berada di media sosial Twitter berupa postingan *tweet* untuk mengetahui sentimen dari opini tersebut terhadap aplikasi Riliv. Sentimen yang dimaksud meliputi sentimen positif, negatif dan netral. Konteks dari masalah berada pada pendapat pengguna Twitter terhadap aplikasi Riliv. Klasifikasi sentimen ini dapat membantu calon pengguna dalam mengetahui apakah aplikasi ini direkomendasikan atau tidak dari perspektif penggunanya.

3.2 Data Acquisition

Proses *data acquisition* dimulai dengan *scraping data* menggunakan *library python sncrape*. Pengumpulan data berupa data *tweet* dari Twitter menggunakan kata kunci 'aplikasi riliv' dalam Bahasa Indonesia selama rentang waktu 1 Januari 2018 hingga 1 Januari 2023, didapatkan data sebanyak 1035 data. Gambar 2 menunjukkan hasil *scraping data*.

1	User	Date Created	Number of Likes	Source of Tweet	Tweet
2	PaduduS	2022-11-26 09:31:30+00:00	0	Twitter Web App	@tanyakanrl coba meditasi pakai aplikasi riliv buat aku sangat membantu"
3	namakuhazelle	2022-02-05 03:49:42+00:00	0	Twitter for iPad	@whatssadness di aplikasi riliv, ada sesi konselingnya. coba ajaa, bagus kokðŸ'
4	jejenyemarkie	2022-07-24 16:29:52+00:00	0	Twitter Web App	aplikasi riliv km kenapa saat aku mau menggunakanmu malah eror..
5	kembangmlathi	2019-10-07 13:21:35+00:00	0	Twitter for iPhone	@kembaranyoungk @AREAJULID @adgilaif Riliv, aplikasi. Mahal tapi.
6	chanxdra	2021-07-27 07:34:56+00:00	1	Twitter for Android	@dohyaun Kalau online coba ke aplikasi riliv juy.
7	nurfajriahk	2022-04-21 14:47:38+00:00	1	Twitter for Android	@cutiesbfp Ada aplikasi namanya riliv, disitu banyak meditasi buat kecemasan.

Gambar 1. Scraping Data

Data *tweet* dari pihak Riliv dengan *user* riliv_app dieliminasi dari *dataset*, sehingga *dataset* awal berjumlah 1035 data berkurang menjadi 621 data.

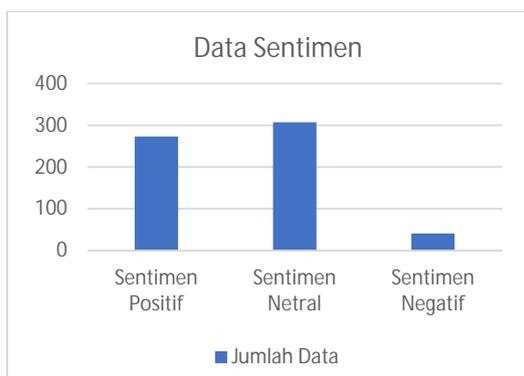
Proses *labelling* secara manual dilakukan dengan label positif, negatif dan netral. Hasil dari proses *labelling* dapat dilihat pada Gambar 3.

1	Tweet	Sentimen
2	@tanyakanrl coba meditasi pakai aplikasi riliv buat aku sangat membantu"	Positif
3	@whatssadness di aplikasi riliv, ada sesi konselingnya. coba ajaa, bagus kokðŸ'	Positif
4	aplikasi riliv km kenapa saat aku mau menggunakanmu malah eror..	Negatif
5	@kembaranyoungk @AREAJULID @adgilaif Riliv, aplikasi. Mahal tapi.	Negatif
6	@dohyaun Kalau online coba ke aplikasi riliv juy.	Netral
7	@cutiesbfp Ada aplikasi namanya riliv, disitu banyak meditasi buat kecemasan.	Netral

Gambar 2. Hasil Labelling

Setelah dilakukan *labelling*, klasifikasi sentimen didapatkan tiga sentimen dengan persebaran yang tertera pada Gambar 4.

pengguna mengetahui tentang aplikasi Riliv dan memberikan ulasan positif mengenai aplikasi Riliv, sedangkan sedikit yang merasa kecewa terhadap aplikasi tersebut.

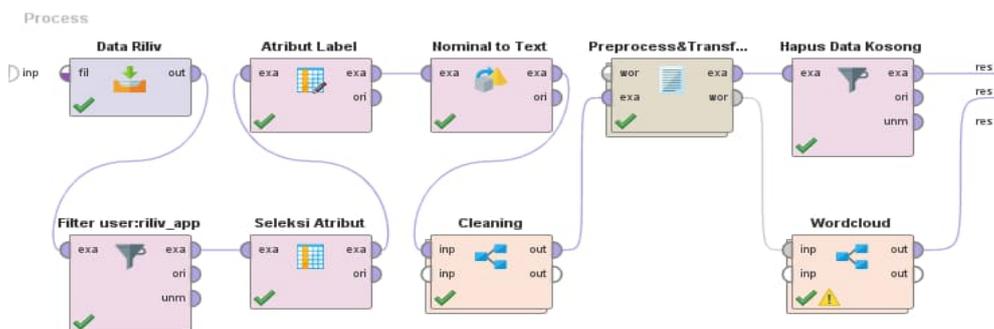


Gambar 3. Klasifikasi Data Sentimen

Terdapat data sentimen sebanyak 273 data positif, 307 data netral dan 41 data negatif. Berdasarkan hasil persebaran data sentimen, sebagian besar

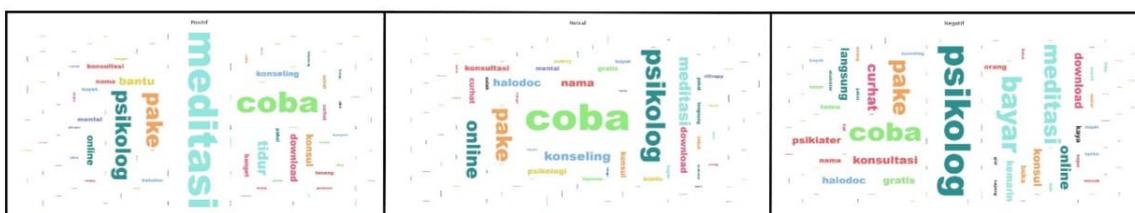
3.3 Data Exploration

Tahap *data exploration* ini terdiri dari tahap *preprocessing*, transformasi dan visualisasi data. Proses pada tahap *data exploration* tertera pada Gambar 5. Operator *cleaning* dan *preprocess&transform* merupakan bagian dari *preprocessing* dan transformasi data. Pada operator *wordcloud* untuk visualisasi data.



Gambar 4. Data exploration di RapidMiner

Penyajian data ditampilkan dibangun berupa seratus kata yang dalam bentuk *wordcloud* dari hasil paling sering muncul dari tiga sentimen. klasifikasi sentimen. Visualisasi yang



Gambar 5. Wordcloud opini positif, netral dan negatif

Kata yang paling sering muncul pada sentimen positif seperti ditunjukkan pada Gambar 6 diantaranya yaitu kata “meditasi”, “coba”, dan “pake”, pada sentimen netral terdapat kata “coba”, “psikolog”, dan “pake”, dan pada sentimen negatif diantaranya yaitu kata “psikolog”, “bayar”, dan “coba”.

3.4 Modelling

Data diproses menerapkan algoritma *Naïve Bayes* dan *Random Forest*. Hasil proses didapatkan data bersih berjumlah 585 data. Kelas minoritas yaitu sentimen negatif disetarakan menggunakan SMOTE, seperti yang tertera pada Tabel 1.

Tabel 1. Hasil SMOTE

Label	Jumlah	
	Sebelum SMOTE	Setelah SMOTE
Positif	273	273
Netral	273	273
Negatif	39	273
Total	585	819

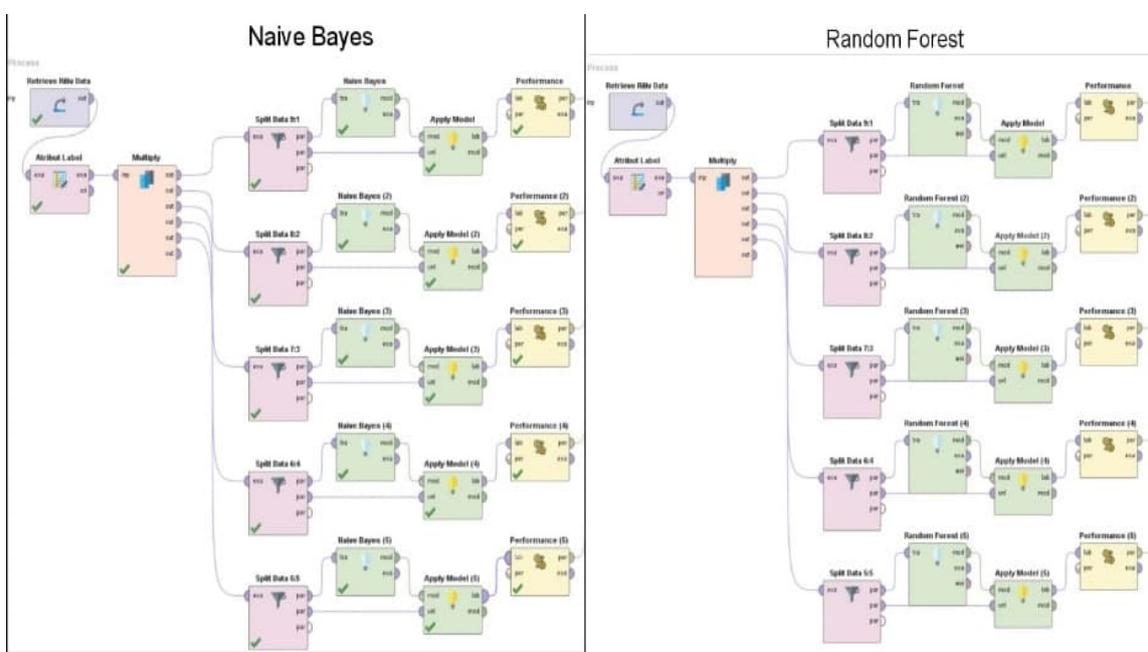
Tahap *modelling* dilakukan proses klasifikasi data, yakni pada *data training* dan *data testing*. Untuk hasil klasifikasi dilakukan perbandingan antara penggunaan algoritma klasifikasi dengan data hasil SMOTE dan tanpa SMOTE. Tabel 2 menunjukkan lima skenario dari pembagian *data training* dan *data testing* untuk klasifikasi yang dilakukan.

Tabel 2. Pembagian *data training* dan *data testing*

Skenario	Split Data	Tanpa SMOTE		Setelah SMOTE	
		Data Training	Data Testing	Data Training	Data Testing
1	90% : 10%	527	58	737	82
2	80% : 20%	468	117	655	164
3	70% : 30%	409	176	573	246
4	60% : 40%	351	234	491	328
5	50% : 50%	293	292	410	409

Pada Gambar 7, terlihat dua proses dari model yang digunakan untuk mengolah data di RapidMiner. Data digandakan menggunakan operator *multiply* supaya dapat diproses sesuai

dengan pembagian *data training* dan *data testing* sesuai Tabel 2. Perbedaan proses dari kedua model terletak pada operator model yang memproses *data training*.



Gambar 6. Model *Naive Bayes* dan *Random Forest*

3.5 Evaluation

Hasil dari klasifikasi menggunakan *Naive Bayes* dengan SMOTE dan tanpa SMOTE menghasilkan nilai evaluasi yang dapat dilihat pada Tabel 3. Nilai evaluasi berupa nilai akurasi, presisi dan *recall*. Hasil evaluasi dari lima skenario tersebut

menunjukkan bahwa evaluasi terbaik dari kedua klasifikasi menggunakan *Naive Bayes* yaitu pada skenario satu dimana persentase pembagian data 90% *data training* dan 10% *data testing* yakni dengan nilai akurasi, presisi dan *recall* yang tertinggi.

Tabel 3. Hasil Evaluasi *Naïve Bayes*

Pembagian Data	<i>Naïve Bayes</i>			<i>Naïve Bayes</i> + SMOTE		
	Akurasi	Presisi	<i>Recall</i>	Akurasi	Presisi	<i>Recall</i>
Skenario 1	72.41%	59.15%	58.95%	82.72%	82.89%	82.72%
Skenario 2	60.17%	47.93%	46.59%	73.94%	74.41%	73.94%
Skenario 3	56.25%	44.92%	42.62%	71.54%	71.45%	71.54%
Skenario 4	50.43%	39.13%	36.09%	66.97%	65.95%	66.97%
Skenario 5	48.63%	39.20%	39.23%	64.79%	63.64%	64.85%

Hasil dari klasifikasi menggunakan *Random Forest* dengan SMOTE dan tanpa SMOTE menghasilkan nilai evaluasi yang dapat dilihat pada tabel 4. Hasil dari lima skenario pada *Random Forest* menunjukkan bahwa evaluasi terbaik

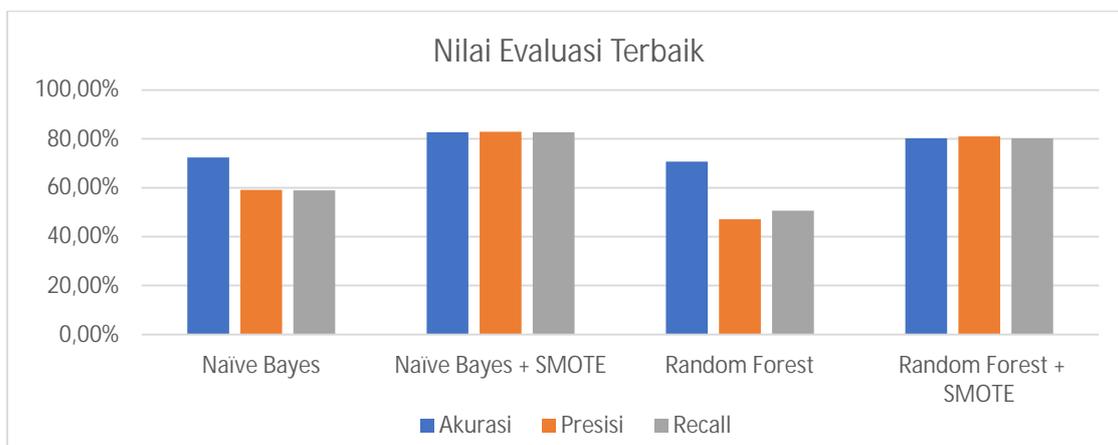
dari kedua klasifikasi *Random Forest* yaitu pada skenario satu dimana persentase pembagian data 90% data *training* dan 10% data *testing* yakni dengan nilai akurasi, presisi dan *recall* yang tertinggi.

Tabel 4. Hasil Evaluasi *Random Forest*

Pembagian Data	<i>Random Forest</i>			<i>Random Forest</i> + SMOTE		
	Akurasi	Presisi	<i>Recall</i>	Akurasi	Presisi	<i>Recall</i>
Skenario 1	70.69%	47.11%	50.62%	80.25%	81.03%	80.25%
Skenario 2	57.63%	40.26%	41.21%	75.76%	75.58%	75.76%
Skenario 3	68.18%	47.09%	48.78%	77.24%	78.22%	77.24%
Skenario 4	61.11%	41.42%	43.73%	75.23%	75.14%	75.23%
Skenario 5	59.59%	39.73%	42.49%	74.33%	74.93%	74.33%

Gambar 8 menampilkan nilai evaluasi terbaik dari klasifikasi *Naïve Bayes* dan *Random Forest* berdasarkan lima skenario pembagian data

yang telah ditentukan, yakni didapatkan pada skenario satu dengan pembagian data *training* 90% dan data *testing* 10%.



Gambar 7. Perbandingan Nilai Evaluasi Model Klasifikasi *Naïve Bayes* dan *Random Forest*

Hasil akurasi pada klasifikasi *Naïve Bayes* dan *Random Forest* mendapatkan keseluruhan evaluasi tertinggi pada data hasil SMOTE. yaitu dengan akurasi sebesar 82.72% dan 80.25%. Evaluasi tertinggi yaitu pada hasil klasifikasi *Naïve Bayes* dengan SMOTE. Evaluasi terbaik kedua yakni hasil klasifikasi *Random Forest* dengan SMOTE, urutan ketiga yakni hasil klasifikasi *Naïve Bayes* dan keempat yaitu hasil klasifikasi *Random Forest*.

4. KESIMPULAN

Hasil performa model dengan *confusion matrix* menghasilkan nilai akurasi terbaik pada model *Naïve Bayes* dan *Random Forest* dengan SMOTE di skenario 1 yakni, pada model *Naïve Bayes* menghasilkan nilai akurasi sebesar 82.72%, *precision* sebesar 82.89%, *recall* sebesar 82.72% dan pada model *Random Forest* menghasilkan nilai akurasi sebesar 80.25%, *precision* sebesar 81.03%, *recall* sebesar 80.25%. Hasil dari kedua algoritma tersebut menunjukkan hasil performa model *Naïve Bayes* yang lebih optimal.

DAFTAR PUSTAKA

Aldean, M. Y., Paradise, P., & Setya Nugraha, N. A. (2022). Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 di Twitter

Menggunakan Metode Random Forest Classifier (Studi Kasus: Vaksin Sinovac). *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, 4(2), 64–72.

<https://doi.org/10.20895/inista.v4i2.575>

Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*. <https://doi.org/10.1016/j.knosys.2021.107134>

Fadhillah, R. P. (2022). *Analisis Sentimen Terhadap Pemberitaan Varian Omicron di Indonesia pada Media Sosial Instagram Menggunakan Naïve Bayes*. Universitas Singaperbangsa Karawang.

Giovani, A. P., Ardiansyah, A., Haryanti, T., Kurniawati, L., & Gata, W. (2020). Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi. *Jurnal Teknoinfo*. <https://doi.org/10.33365/jti.v14i2.679>

Halim, C., Purnomo, H. D., & Wahyono, T. (2022). Analisis Pengelompokan Wilayah Penyebaran Covid-19 Di Indonesia Dengan Metode Clustering Menggunakan Algoritma K-Means dan K-Medoids. *Jurnal Inovtek Polbeng*, 7(2), 359–372.

Hendra, A., & Fitriyani, F. (2021). Analisis Sentimen Review Halodoc Menggunakan Naïve Bayes Classifier. *JISKA (Jurnal*

- Informatika Sunan Kalijaga*.
<https://doi.org/10.14421/jiska.2021.6.2.78-89>
- Nanda, S., Mualfah, D., & Fitri, D. (2022). Analisis Sentimen Kepuasan Pengguna Terhadap Layanan Streaming Mola Menggunakan Algoritma Random Forest. *Jurnal Aplikasi Teknologi Informasi Dan Manajemen (JATIM)*, 3(2 SE-Articles).
<https://doi.org/10.31102/jatim.v3i2.1592>
- Qadrini, L., Hikmah, H., & Megasari, M. (2022). Oversampling, Undersampling, Smote SVM dan Random Forest pada Klasifikasi Penerima Bidikmisi Sejava Timur Tahun 2017. *Journal of Computer System and Informatics (JoSYC)*, 3(4), 386–391.
<https://doi.org/10.47065/josyc.v3i4.2154>
- Saputra, S. A., Didi Rosiyadi, Windu Gata, & Syepri Maulana Husain. (2019). Sentiment Analysis Analysis of E-Wallet Sentiments on Google Play Using the Naive Bayes Algorithm Based on Particle Swarm Optimization. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 3(3), 377–382.
<https://doi.org/10.29207/resti.v3i3.1118>
- Tiara, O. R. (2021). *Peduli Kesehatan Mental dengan Startup Asal Indonesia Ini*. My Skill.
<https://myskill.id/blog/dunia-kerja/kesehatan-mental-startup-indonesia/>
- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*.
<https://doi.org/10.1007/s10462-022-10144-1>
- Wardani, N. R., & Erfina, A. (2021). Analisis Sentimen Masyarakat Terhadap Layanan Konsultasi Dokter Menggunakan Algoritma Naive Bayes. *SISMATIK (Seminar Nasional Sistem Informasi Dan Manajemen Informatika)*, 11–18.
<https://sismatik.nusaputra.ac.id/>
- Wijoto, R. (2019). *Aplikasi Karya Anak Surabaya Ini Kantongi Penghargaan dari Google*. Beritajatim.Com.
<https://beritajatim.com/teknologi/aplikasi-karya-anak-surabaya-ini-kantongi-penghargaan-dari-google/>