

## ANALISIS SENTIMEN PEMINDAHAN IBU KOTA NEGARA (IKN) MENGGUNAKAN METODE OVERSAMPLING SYNTHETIC MINORITY (SMOTE)

AHMAD ALMAS'UD ZD<sup>1</sup>, KUSRINI<sup>2</sup>, ANGGIT DWI HARTANTO<sup>3</sup>

<sup>1,2,3</sup>Magister Teknik Informatika, Universitas Amikom Yogyakarta

Email : [ahmadalmasudzd@students.amikom.ac.id](mailto:ahmadalmasudzd@students.amikom.ac.id)<sup>1</sup>, [kusrini@amikom.ac.id](mailto:kusrini@amikom.ac.id)<sup>2</sup>, [anggit@amikom.ac.id](mailto:anggit@amikom.ac.id)<sup>3</sup>

### ABSTRAK

Pemindahan Ibu Kota Negara (IKN) adalah keputusan strategis yang memicu berbagai tanggapan dan reaksi dari berbagai pihak. Analisis sentimen terhadap pemindahan IKN menjadi suatu aspek penting untuk memahami pola pikir dan sikap masyarakat terhadap keputusan tersebut. Studi ini bertujuan untuk melakukan perbandingan performa algoritma klasifikasi terhadap analisis sentimen pada data teks yang berisi opini dan pendapat masyarakat terkait pemindahan IKN menggunakan metode *oversampling Synthetic Minority (SMOTE)* untuk mengatasi masalah ketidakseimbangan dalam data sentimen. Dua algoritma klasifikasi, yaitu *Support Vector Machine (SVM)* dan *Random Forest*, dievaluasi dalam konteks ini. Hasil penelitian menunjukkan peningkatan signifikan dalam performa kedua algoritma setelah penerapan metode SMOTE. Performa algoritma SVM meningkat dari 85% menjadi 92%, sementara algoritma *Random Forest* meningkat dari 84% menjadi 91%. Hasil ini menunjukkan bahwa metode SMOTE efektif dalam meningkatkan kemampuan algoritma klasifikasi untuk mengatasi ketidakseimbangan data dan menghasilkan prediksi sentimen yang lebih akurat.

**Kata kunci :** IKN, *Random Forest*, *SMOTE*, *Support Vector Machine*.

### I. PENDAHULUAN

Sejak tahun 2019 bangsa Indonesia diramaikan dengan berita tentang pemindahan Ibu Kota Negara (IKN). Meski rencana pemindahan IKN bukanlah rencana baru, mengingat hal ini pernah direncanakan oleh presiden pertama republik Indonesia bapak ir. Soekarno pada tanggal 17 Juli 1957 dengan memilih

palangkaraya sebagai ibu kota negara. Pada masa orde baru, yakni tahun 1990-an, presiden kedua republik Indonesia juga merencanakan pemindahan IKN ke Jonggol Jawa Barat. Namun, pemindahan IKN baru serius digarap pada masa pemerintahan presiden Ir. H. Joko Widodo, tepatnya pada tahun 2017 dimana upaya tersebut kembali dimunculkan oleh kementrian. Pembangunan Ibu Kota Nusantara pada pertengahan Maret 2022 akan direlokasi secara bertahap dari tahun 2024 hingga 2045. Tepat pada Senin, 26 Agustus 2019, Melalui siaran pers, lokasi baru IKN berada di Kabupaten Penajam Paser Utara Provinsi Kalimantan Timur, tepatnya di sebagian Kabupaten Kutai Kartanegara dan disahkannya RUU IKN menjadi UU IKN pada tanggal 18 Januari 2022. Rencana pemerintah memindahkan IKN beralaskan pada visi misi Indonesia tahun 2045 yaitu Indonesia maju, ekonomi Indonesia akan masuk 5 besar dunia 2045 (Nainggolan, 2022), pemindahan IKN akan mendorong pertumbuhan ekonomi yang merata termasuk di kawasan timur Indonesia (Aziz, 2022) Namun demikian pemindahan IKN juga akan beresiko pada tingginya pembiayaan pemindahan ibukota, selain itu adanya pemindahan ibu kota negara Indonesia, tentu mengundang berbagai reaksi, terutama bagi masyarakat Indonesia, mengingat ibu kota negara baru di Indonesia merupakan hal yang sensitif sehingga ramai diperbincangkan di media sosial termasuk Twitter.

*Twitter* adalah sebuah *platform microblogging* yang memungkinkan penggunaanya mengirimkan serta memperbaharui pesan singkat secara *real time*. Jenis pesan singkat yang dikirimkan pada aplikasi twitter lebih dikenal dengan sebutan *tweet*. *Tweet* merupakan pesan singkat yang memiliki batasan panjang dengan 280 karakter, karena keterbatasan karakter itulah yang menyebabkan penggunaanya seringkali menggunakan kata singkatan, Bahasa *slang*, maupun kesalahan pengejaan (Nurrun Muchammad Shiddieqy et al., 2016). *Twitter* memuat banyak sekali opini masyarakat terkait suatu topik pembahasan. Opini yang terdapat pada *twitter* cukup variatif, hal tersebut dikarenakan pengguna *twitter* memiliki latar belakang dan berasal dari golongan masyarakat yang berbeda. Dengan demikian *tweet* tersebut dapat diolah menjadi informasi yang bermanfaat dengan teknik analisis sentimen.

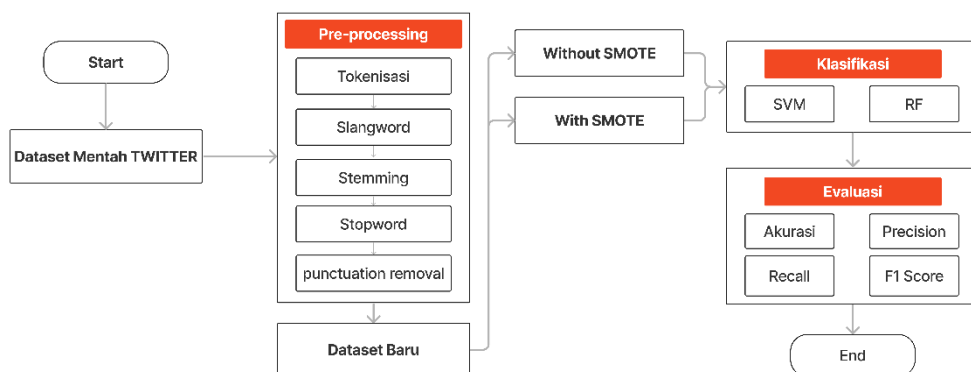
Analisis sentimen adalah sebuah tahapan memahami, mengekstraksi, dan

secara otomatis mendapatkan sebuah informasi (Cindo et al., 2019). Analisis sentimen terhadap komentar pada laman media sosial *twitter*, memberikan pandangan berharga terhadap persepsi dan pandangan masyarakat terkait pemindahan Ibu Kota Negara.

Oleh karena itu, Penelitian ini bertujuan untuk melakukan pengujian performa antara algoritma *Support Vector Machine (SVM)* dan *Random Forest* berdasarkan analisis sentimen terhadap komentar di laman media sosial *twitter* terkait pemindahan Ibu Kota Negara., dengan menggunakan metode SMOTE sebagai alat utama. Dalam hal ini, metode Oversampling Synthetic Minority (SMOTE) memainkan peran penting dalam meningkatkan validitas analisis (Barro et al., 2013). SMOTE, sebagai teknik oversampling yang fokus pada kelas minoritas, memungkinkan penanganan seimbang terhadap opini yang mungkin kurang terwakili dalam data komentar.

## II. METODE PENELITIAN

Pada penelitian ini metode klasifikasi yang digunakan adalah *Support Vector Machine (SVM)* dan *Random Forest*. Untuk mendapatkan hasil terbaik, pada penelitian ini akan dilakukan berbagai tahapan yang meliputi pengumpulan data, *pre-processing* data, lalu dilanjutkan dengan teknik *oversampling method* kemudian diimplementasikan metode klasifikasi menggunakan algoritma *Support Vector Machine (SVM)* dan *Random Forest*. Tahapan tersebut dapat dilihat pada gambar 1.



Gambar 1. Alur Penelitian

### **A. Metode Pengumpulan Data**

Metode pengumpulan data merupakan tahapan awal pada sebuah penelitian. Dalam penelitian ini, data yang dikumpulkan berasal dari berbagai komentar pada platform *twitter* yang berkaitan dengan pemindahan Ibu Kota Negara.

### **B. Pre-processing**

*Pre-processing* adalah tahapan awal untuk mempersiapkan teks menjadi data yang dapat di proses lebih lanjut, selain itu *pre-processing* ini bertujuan untuk mempersiapkan data sebelum diproses di tahapan ekstraksi kata kunci dan menentukan sentimen (Amalia, 2023). Untuk memaksimalkan tahapan *pre-processing* lebih maksimal, penelitian ini melakukan beberapa tahapan *pre-processing* yaitu :

1. Tokenisasi : proses ini menghilangkan kata-kata yang tidak penting atau kata-kata yang tidak mengandung opini.
2. *Case-folding* : proses yang menjadikan setiap huruf besar ke huruf kecil agar tidak terjadi kesalahan pencocokan karakter atau huruf dalam suatu kata
3. *Stemming* : Proses stemming adalah proses pemetaan dan penguraian berbagai bentuk (varian) menjadi bentuk-bentuk stem (kata dasar) berdasarkan kata yang dihasilkan dengan menghilangkan stopword atau filtering. Proses stemming digunakan untuk meningkatkan kualitas informasi dari data yang digunakan. Salah satu cara untuk melakukan proses stemming yaitu menggunakan metode kamus. Misalnya, tiga kata “bekerja”, “bekerja” dan “pekerja” yang semula memiliki arti yang berbeda dapat diekstraksi menjadi kata “bekerja” dengan arti yang sama, sehingga membuat kata- kata tersebut saling berhubungan
4. *Stopword removal* : proses ini menghilangkan kata-kata yang tidak penting atau kata-kata yang tidak mengandung opini.
5. *Punctuation removal* : proses yang memisahkan tanda baca dan

simbol lainnya selain alfabet, yang tujuannya agar tanda baca dan simbol-simbol tersebut tidak terproses saat proses dan tahap selanjutnya (Undap et al., 2021).

### C. *Oversampling Synthetic Minority (SMOTE)*

Pada metode SMOTE prinsip yang digunakan adalah menambah jumlah data kelas minor agar setara dengan kelas mayor dengan cara membangkitkan synthesis data atau data buatan (Sofyan & Prasetyo, 2021). Sebuah teknik oversampling yang digunakan dalam pembelajaran mesin untuk menangani masalah *imbalanced class* pada data. Metode SMOTE bekerja dengan menambahkan jumlah data pada kelas minor hingga sama dengan kelas mayor dengan cara membuat data sintetis. Data sintetis tersebut dibuat berdasarkan nilai k-tetangga terdekat dari kelas minor. Prinsip dari SMOTE adalah untuk menyeimbangkan data kelas dengan cara membangkitkan data sintetis (Nugroho & Rilvani, 2023).

Cara kerja SMOTE diasumsikan pada data dengan jumlah variabel p maka jarak antara  $x^T = [x_1, x_2, \dots, x_p]$  dan  $z^T = [z_1, z_2, \dots, z_p]$  adalah  $d(x, y) = \sqrt{(x_1 - z_1)^2 + (x_2 - z_2)^2 + \dots + (x_p - z_p)^2}$ . Maka untuk memunculkan data dengan pendekatan SMOTE digunakan persamaan sebagai berikut :

$$x_{syn} = x_i + (x_{knn} - x_i)\gamma$$

$x_{syn}$  merupakan pengamatan baru hasil pembangkitan,  $x_i$  adalah pengamatan ke- $i$ ,  $x_{knn}$  merupakan  $x$  terdekat dari  $x_i$ , serta  $\gamma$  merupakan bilangan acak antara 0 dan 1. Untuk data nominal maka akan diisi dengan nilai mayoritas pada k-tetangga terdekat. Perhitungan jarak pada SMOTE apabila ada variabel kategorik maka akan diganti dengan kuadrat median standar deviasi variabel kontinyu kelas minoritas jika nilai kategorik pada pengamatan ke- $i$  dan  $j$  berbeda (Syukron et al., 2020).

Dapat diasumsikan SMOTE bekerja dengan mengidentifikasi sampel-sampel minoritas dalam dataset, yaitu sampel-sampel yang berasal dari kelas yang kurang jumlahnya antara lain (1) Setiap sampel minoritas dipilih beberapa tetangga terdekat dari kelas yang sama. Jumlah tetangga yang dipilih menjadi faktor penting dalam menghasilkan sampel sintetis. (2) SMOTE kemudian

membuat sampel sintetis dengan cara menggabungkan informasi dari sampel minoritas yang dipilih dan tetangga terdekat. Ini dilakukan dengan menghitung perbedaan antara atribut-atribut sampel dan tetangga, kemudian mengalikan perbedaan tersebut dengan suatu faktor dan menambahkannya ke sampel minoritas awal. (3) Sampel-sampel sintetis yang dihasilkan kemudian ditambahkan ke dataset asli untuk membentuk dataset yang baru. Dengan penambahan ini, jumlah sampel kelas minoritas meningkat, dan kelas minoritas menjadi lebih seimbang dengan kelas mayoritas.

#### **D. Klasifikasi**

Dalam proses pembelajaran mesin, klasifikasi adalah tugas untuk memprediksi label kelas dari sampel yang diberikan berdasarkan sekumpulan fitur atau karakteristik.

Pada tahapan ini, digunakan 2 model klasifikasi yaitu :

##### **1. Support Vector Machine (SVM)**

Algoritma *Support Vector Machine (SVM)* merupakan bagian dari metode klasifikasi pada *data mining* yang bisa melakukan permasalahan linear dan non-linear sehingga dapat digunakan dalam metode klasifikasi data khususnya pada data mining. Menurut Santoso dalam Bitu Praga Zen (Zen et al., 2022) Kelebihan dari algoritma *Support Vector Machine (SVM)* adalah dari kemampuannya untuk menerapkan pemisah linear pada input data non-linear berdimensi tinggi, dan ini diperoleh dengan menggunakan fungsi kernel yang dipilih dan diterapkan berdasarkan karakteristik data. Banyak peneliti telah melaporkan bahwa *Support Vector Machine (SVM)* metode yang paling akurat untuk teks klasifikasi. Dalam pemodelan klasifikasi *Support Vector Machine (SVM)* memiliki konsep yang lebih matang dan lebih jelas secara matematis dibandingkan dengan teknik-teknik klasifikasi lainnya. *Support Vector Machine (SVM)* digunakan untuk mencari *hyperlane* terbaik dengan memaksimalkan jarak antar kelas. *Hyperlane* adalah sebuah fungsi yang dapat digunakan untuk pemisah antar kelas.

**2. Random Forest**

Random Forest merupakan algoritma dalam machine learning yang digunakan untuk pengklasifikasian dataset dalam jumlah besar. Karena fungsinya bisa digunakan untuk banyak dimensi dengan berbagai skala dan performa yang tinggi. Random Forest diimplementasikan untuk berbagai sektor seperti perbankan, kesehatan, analisis keuangan, ecommerce dan sebagainya. Cara kerja algoritma ini adalah dengan membangun beberapa pohon keputusan dan menggabungkannya untuk mendapatkan prediksi yang lebih stabil dan akurat. Hutan (Forest) yang dibangun oleh algoritma tersebut merupakan kumpulan pohon keputusan dimana biasanya dilatih dengan metode bagging, yang diharapkan menghasilkan model machine learning yang lebih baik (Sudrajat et al., 2022).

**E. Hasil dan Evaluasi**

Pada tahapan akhir ini akan dilakukan beberapa parameter uji seperti *accuracy*, *f1 score*, *recall* dan *precision*.

**III. HASIL DAN PEMBAHASAN**

**A. Dataset mentah twitter**

Berdasarkan dataset yang didapatkan pada *platform twitter*, didapatkan dataset mentah berjumlah 1.506 baris *tweet*.

Tabel 1. Dataset Mentah *Twitter*

Index	tweet	username
0	Dulu ad seorang bilang begini, dibilangnya benci sama kalimantan, sekarang rasain aja deh, gue kagak peduli, lha yang anti IKN itu tujuannya baik, agar tdk ada suku yg terusir 🤔	Gusuran76
1	@SahronihO @machbeach 🤔 Saran saya kalau punya OTAK peke ! Bodoh memang pilihan , tapi TOLOL jgn bangga ! Bagaimana dana haji mau di pakai IKN Laah negara saja masih memberikan SUBSIDI untuk biaya haji ! 🤔 Berita hoax di pakai referensi Ketololan berfikir 🤔 Otak kamu WARAS @SahronihO ?	syahidah_d
2	Pemerintah resmi membentuk TIM Transisi Pemindahan IKN . 🤔🤔🤔🤔IDID	mimakhayla

3	Muncul Petisi Sejumlah Tokoh Penting Tolak IKN, Ruhut: Enggak Usah Cari Popularitas Murahahan - Terimakasih Mudik Lancar	Mah_Mood2
4	@Sangkuriang5551 Waduh ... Jadi sumber dana IKN bukan saja dari "sumbangan sukarela masyarakat", namun dpt dibilang "iuran wajib dari masyarakat" (atau identik dg 'pemalakan" ? ) #IndonesiaDaruratUtang #IndonesiaDaruratUtang	EdiKecepat

## B. Pre-processing

Pada tahapan ini, dataset mentah *twitter* yang sebelumnya telah didapatkan, dilakukan proses pembersihan kata menggunakan metode *tokenize*, *case-folding*, *stemming*, *stopword removal*, dan *punctuation removal*.

Tabel 2. Dataset setelah dilakukan Pre-processing

Index	tweet	username
0	dulu ada seorang bilang begini dibilangnya benci sama kalimantan sekarang rasain saja deh saya tidak peduli lha yang anti ikn itu tujuannya baik agar tidak ada suku yang terusir	Gusuran76
1	saran saya kalau punya otak peke bodoh memang pilihan tapi tolong jangan bangga bagaimana dana haji mau di pakai ikn laah negara saja masih memberikan subsidi untuk biaya haji berita isu bohong di pakai referensi ketololan berfikir otak kamu waras	syahidah_d
2	pemerintah resmi membentuk tim transisi pemindahan ikn strategi wujudkannkrimaju	mimakhayla
3	muncul petisi sejumlah tokoh penting tolak ikn ruhut tidak perlu cari popularitas murahahan terimakasih mudik lancar	Mah_Mood2
4	waduh jadi sumber dana ikn bukan saja dari sumbangan sukarela masyarakat namun dapat dibilang iuran wajib dari masyarakat atau identik dengan pemalakan indonesiaaruratutang indonesiaaruratutang	EdiKecepat

Setelah dilakukan proses *pre-processing* data menggunakan *case-folding* dan *cleaning* selanjutnya dilakukan proses pelabelan data menggunakan *library TextBlob*. *TextBlob* merupakan sebuah *library python* yang berfungsi untuk melakukan proses analisis kata (NLP) salah satunya pada tahapan analisis sentimen.

Tabel 3. Dataset setelah proses labeling

Index	tweet	label
0	dulu ada seorang bilang begini dibilangnya benci sama kalimantan sekarang rasain saja deh saya tidak peduli lha yang anti ikn itu tujuannya baik agar tidak ada suku yang terusir	negatif
1	saran saya kalau punya otak peke bodoh memang pilihan tapi tolong jangan bangga bagaimana dana haji mau di pakai ikn laah negara saja masih memberikan subsidi untuk biaya haji berita isu bohong di pakai referensi ketololan berfikir otak kamu waras	negatif



2	pemerintah resmi membentuk tim transisi pemindahan ikn strategi wujudkan krimaju	negatif
3	muncul petisi sejumlah tokoh penting tolak ikn ruhut tidak perlu cari popularitas murahan terimakasih mudik lancar	negatif
4	waduh jadi sumber dana ikn bukan saja dari sumbangan sukarela masyarakat namun dapat dibilang iuran wajib dari masyarakat atau identik dengan pemalakan indonesia darurat utang indonesia darurat utang	negatif

Setelah dilakukan proses pre-processing dan preparation data , data berkurang menjadi 1.030 data sebagai berikut :

Tabel 4. Jumlah dataset

Kelas	Jumlah Dataset
Positif	324
Negatif	706
<b>Total</b>	<b>1030</b>

### C. Implementasi SMOTE

Sebelum melakukan proses SMOTE, dilakukan tahap *pre-processing* dan preparation data. Langkah ini melibatkan normalisasi, penghapusan nilai yang hilang, atau transformasi data agar dapat digunakan secara efektif dalam proses selanjutnya. Setelah tahap *pre-processing*, langkah berikutnya adalah menangani ketidakseimbangan kelas pada dataset. Pada penelitian ini digunakan metode SMOTE untuk menyeimbangkan jumlah dataset antara kelas positif dan negatif. SMOTE bekerja dengan cara menciptakan data sintetis pada kelas minoritas (kelas positif) untuk menyamakan jumlahnya dengan kelas mayoritas (kelas negatif). Proses SMOTE dilakukan dengan memilih sampel-sampel dari kelas positif dan mengidentifikasi tetangga terdekatnya. Kemudian, sampel sintetis dibuat dengan menggabungkan atribut-atribut dari sampel minoritas dan tetangga terdekat dengan faktor tertentu. Pada akhirnya, sampel sintetis ini ditambahkan ke dalam dataset.

Tabel 5. Hasil Implementasi SMOTE

Kelas	Jumlah Dataset
Positif	706
Negatif	706

Tabel 5 menunjukkan hasil dari implementasi SMOTE, yang mengindikasikan bahwa kelas positif dan negatif kini memiliki jumlah dataset yang sama, yaitu sebanyak 706. Jumlah ini mencerminkan keberhasilan SMOTE dalam meningkatkan jumlah sampel pada kelas positif dari 324 menjadi 706. Hal ini dikarenakan metode SMOTE bekerja dengan cara menambahkan jumlah data pada kelas sentimen positif hingga sama dengan kelas bersentimen negatif dengan cara membuat data sintetis, sehingga data pada kelas positif dan negatif berjumlah sama. Penerapan SMOTE memiliki dampak signifikan dalam meningkatkan keseimbangan dataset, memastikan bahwa model yang akan dikembangkan tidak akan menjadi bias terhadap kelas mayoritas. Keseimbangan yang dicapai dengan SMOTE dapat meningkatkan kinerja model dalam mengenali sentimen positif, karena model memiliki lebih banyak data untuk dilatih pada kelas minoritas (Erlin, E, dkk, 2022).

#### D. Hasil dan Evaluasi

Setelah dilakukan proses *balancing* data menggunakan algoritma SMOTE, maka pada tahapan ini, dilakukan kombinasi antara algoritma *Support Vector Machine (SVM)* + SMOTE dan *Random Forest* + SMOTE. Berdasarkan tahapan tersebut, maka didapatkan hasil sebagai berikut :

Tabel 6. Hasil SVM

Kelas	Accuracy	Recall	F1-Score	Precision
	<b>0.85</b>	-	-	-
Positif	-	0.58	0.72	0.92
Negatif	-	0.98	0.89	0.82

Tabel 7. Hasil SVM + SMOTE

Kelas	Accuracy	Recall	F1-Score	Precision
	<b>0.92</b>	-	-	-
Positif	-	0.86	0.92	0.97
Negatif	-	0.98	0.91	0.86

Pada tabel 6 dan 7 diatas, memperlihatkan bahwa nilai akurasi pada algoritma *Support Vector Machine (SVM)* sebelum diterapkan metode SMOTE dan sesudah diterapkan metode SMOTE hasilnya terjadi peningkatan nilai akurasi yang sangat signifikan yaitu bertambah menjadi 0.17

Tabel 8. Hasil Random Forest

Kelas	Accuracy	Recall	F1-Score	Precision
	<b>0.84</b>	-	-	-
Positif	-	0.69	0.74	0.80
Negatif	-	0.91	0.88	0.85

Tabel 8. Hasil Random Forest + SMOTE

Kelas	Accuracy	Recall	F1-Score	Precision
	<b>0.91</b>	-	-	-
Positif	-	0.90	0.91	0.92
Negatif	-	0.91	0.91	0.90

Pada tabel 8 dan 9 diatas, memperlihatkan bahwa nilai akurasi pada algoritma *Random Forest* sebelum diterapkan metode SMOTE dan sesudah diterapkan metode SMOTE hasilnya terjadi peningkatan nilai akurasi yang sangat signifikan yaitu bertambah menjadi 0.17

#### IV. KESIMPULAN

Dengan diterapkannya metode SMOTE kedalam masing-masing algoritma membuat peningkatan performa yang cukup signifikan dengan meningkatnya performa algoritma *Support Vector Machine (SVM)* yang semula bernilai 85% menjadi 92% dan algoritma *Random Forest* yang semula bernilai 84% menjadi 91%. Hasil ini menunjukkan bahwa metode SMOTE efektif dalam meningkatkan kemampuan algoritma klasifikasi untuk mengatasi ketidakseimbangan data dan menghasilkan prediksi sentimen yang lebih akurat.

### DAFTAR PUSTAKA

- Amalia, N. L. (2023). *Analisis Sentimen Pada Perpindahan Ibukota Indonesia Dengan Algoritma Support Vector Machine : Evaluasi Leksikon , Metode Ekstraksi Fitur , Dan Kernel Trick*.
- Barro, R. A., Sulvianti, I. D., & Afendi, F. M. (2013). Penerapan Synthetic Minority Oversampling Technique (SMOTE) Terhadap Data Tidak Seimbang Pada Pembuatan Model Komposisi Jamu. *Journal of Statistics*, 1(1), 1–6. <https://doi.org/10.29244/xplore.v1i1.12424>
- Cindo, M., Dian Palupi Rini, & Ermatita. (2019). Studi Komparatif Metode Ekstraksi Fitur pada Analisis Sentimen Maskapai Penerbangan Menggunakan Support Vector Machine dan Maximum Entropy. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 3(3), 402–407. <https://doi.org/10.29207/resti.v3i3.1159>
- Erlin, E., Desnelita, Y., Nasution, N., Suryati, L., & Zoromi, F. (2022). Dampak SMOTE terhadap Kinerja Random Forest Classifier berdasarkan Data Tidak seimbang. *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, 21(3), 677-690.
- Nugroho, A., & Rilvani, E. (2023). Penerapan Metode Oversampling SMOTE Pada Algoritma Random Forest Untuk Prediksi Kebangkrutan Perusahaan. *Techno.Com*, 22(1), 207–214. <https://doi.org/10.33633/tc.v22i1.7527>
- Nurrun Muchammad Shiddieqy, H., Paulus Insap, S., & Wing Wahyu, W. (2016). Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter. *Seminar Nasional Teknologi Informasi Dan Komunikasi*, 2016(March), 57–64.
- Sofyan, S., & Prasetyo, A. (2021, November). Penerapan Synthetic Minority Oversampling Technique (SMOTE) Terhadap Data Tidak Seimbang Pada Tingkat Pendapatan Pekerja Informal Di Provinsi DI Yogyakarta Tahun 2019. In *Seminar Nasional Official Statistics* (Vol. 2021, No. 1, pp. 868-877).
- Sudrajat, D., Purnamasari, A. I., Dikananda, A. R., Kurnia, D. A., & Bahtiar, A. (2022). Klasifikasi Mutu Pembelajaran Hybrid berdasarkan Algoritma C.45, Random Forest dan Naïve Bayes dengan Optimasi Bootsrap Areggating (Bagging) pada masa COVID-19. *JURIKOM (Jurnal Riset Komputer)*, 9(6), 2227. <https://doi.org/10.30865/jurikom.v9i6.5179>
- Syukron, M., Santoso, R., & Widiharih, T. (2020). Perbandingan Metode Smote Random Forest Dan Smote Xgboost Untuk Klasifikasi Tingkat Penyakit Hepatitis C Pada Imbalance Class Data. *Jurnal Gaussian*, 9(3), 227–236. <https://doi.org/10.14710/j.gauss.v9i3.28915>
- Undap, M. G., Rantung, V. P., & Rompas, P. T. D. (2021). *Analisis Sentimen Situs Pembajak Artikel Penelitian Menggunakan Metode Lexicon-Based*.
- Zen, B. P., Wicaksana, D., & Alfidzar, H. (2022). Analisis Sentimen Tweet Vaksin Covid 19 Sinovac Menggunakan Metode Support Vecor Machine. *Jdmsi*, 3(2), 21–27.