

Analisis Sentimen dengan Metode *Naïve Bayes*, *SMOTE* dan *Adaboost* pada *Twitter* Bank BTN

Kurnia ^{1*}, Indah Purnamasari ², Dedi Dwi Saputra ³

^{1,2,3} Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Nusa Mandiri, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta, Indonesia.

article info

Article history:

Received 10 July 2022

Received in revised form

2 December 2022

Accepted 16 February 2023

Available online April 2023

DOI:

<https://doi.org/10.35870/jtik.v7i2.707>

Keywords:

Sentiment Analysis; *Naïve Bayes*, *SMOTE*; *Adaboost*; Bank BTN.

Kata Kunci:

Analisa Sentimen; *Naïve Bayes*, *SMOTE*; *Adaboost*; Bank BTN.

abstract

Social media has become a means of sharing information with each other, from being a means of promotion to being used to spread opinions. Because of the ease of technology, anyone can access social media and comment on issues that are being discussed. Branding image is also an important thing, because people can interact directly through social media, be it support or criticism. So this study was conducted to analyze the sentiment on twitter towards Bank BTN as a bank that has long focused on housing loans. To analyze the sentiment, an experiment was carried out by a combination of SMOTE, Naïve Bayes and Adaboost algorithms. Before calculating the algorithm, stemming and stopwords are carried out so that the data used does not contain noise. The results showed that the combination of SMOTE, Naïve Bayes, and Adaboost showed the best modeling results with an accuracy of 87.05%, precision of 90.63%, recall of 83.00%, and AUC of 0.909.

abstract

Media sosial menjadi sarana untuk saling berbagi informasi, dari sebagai sarana promosi hingga digunakan untuk menyebarkan opini. Karena kemudahan teknologi, dimanapun orang dapat mengakses media sosial dan mengomentari isu yang sedang ramai dibicarakan. Branding image juga menjadi suatu hal yang penting, karena masyarakat dapat berinteraksi langsung lewat media sosial baik itu dukungan atau kritikan. Sehingga penelitian ini dilakukan untuk menganalisa sentiment yang ada pada media sosial twitter terhadap Bank BTN sebagai Bank yang sudah lama berfokus dalam bidang Kredit Pemilikan Rumah. Untuk menganalisa sentimen tersebut dilakukan percobaan kombinasi algoritma SMOTE, Naïve Bayes dan Adaboost. Sebelum dilakukan perhitungan algoritma, dilakukan tahapan stemming dan stopwords agar data yang digunakan tidak mengandung noise. Hasil penelitian menunjukkan bahwa kombinasi SMOTE, Naïve Bayes, dan Adaboost menunjukkan hasil pemodelan yang paling baik dengan akurasi sebesar 87.05%, presisi sebesar 90.63%, recall sebesar 83.00%, dan AUC sebesar 0.909.

Author. Email: 11211155@nusamandiri.ac.id ^{1}, indah.ih@nusamandiri.ac.id ², dedi.eis@nusamandiri.ac.id ³.

1. Latar Belakang

Media sosial menjadi sarana untuk saling berbagi informasi bahkan dapat digunakan untuk sarana promosi, baik untuk perusahaan atau produk tertentu menurut Kotler dan Keller dalam Augustinah [1]. Bahkan dalam hal pariwisata, hingga politik. Karena media sosial juga bisa digunakan untuk menyebarkan opini-opini baik itu negatif ataupun positif. Adanya gadget, yang membuat setiap orang dapat menggunakan media sosial dimanapun selama terdapat akses *internet* [2]. Karena hal tersebut, maka setiap orang atau entitas perusahaan berusaha membangun media sosial sebagai jembatan komunikasi. Sehingga menjadi penting untuk sebuah perusahaan untuk membangun suatu image branding yang baik di masyarakat sehingga mendapat umpan balik yang positif.

Diantara media sosial populer saat ini salah satunya adalah *Twitter*, dimana saat ini di Indonesia digunakan sebanyak 58,3 % pengguna media sosial dengan jumlah pengguna sebesar 18,45 Juta orang [3]. Para pengguna tersebut menggunakan *Twitter* sebagai tempat untuk mengunggah atau menyampaikan opini pribadinya ataupun mengomentari suatu isu yang timbul dengan adanya fitur *Trending Topics* untuk mengetahui isu yang paling banyak diperbincangkan. Bagi entitas *brand* atau perusahaan, *twitter* juga digunakan sebagai wadah untuk dapat berinteraksi secara langsung dengan pengguna nya dengan komentar atau respon yang beragam. Bahkan karena *twitter* itu sangat terbuka, tidak jarang pula banyak yang menyampaikan pertanyaan atau keluhannya secara langsung lewat akun-akun media sosial *brand* atau perusahaan ini. Untuk itu peneliti bermaksud untuk melakukan penelitian terhadap respon masyarakat yang ada pada media sosial dan menganalisa sentimen yang ada sehingga dapat menjadi suatu bentuk klasifikasi yang dapat diterapkan.

Penelitian ini menggunakan data dari *Twitter* Bank BTN dengan metode *Crawling* data menggunakan *Rapidminer*. Bank BTN dijadikan objek penelitian karena sejarah panjangnya dan fokus bisnis Bank ini yaitu pada produk Kredit Pemilikan Rumah dan telah melayani masyarakat selama 72 tahun. Tentunya hal ini menimbulkan sentiment di masyarakat atas pelayanan Bank BTN selama ini. Berdasarkan latar

belakang tersebut, maka dilakukan penelitian ini untuk melihat sentiment yang ada kepada Bank BTN dari masyarakat dan dilakukan analisa sentimen dengan klasifikasi *Positive* dan *Negative*.

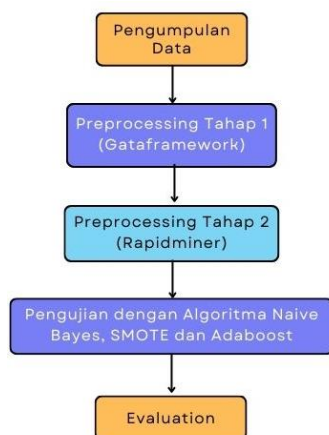
Penelitian ini dilakukan menggunakan algoritma *Naïve Bayes*, *SMOTE* dan *Adaboost* untuk melihat kombinasi algoritma mana yang paling baik dalam melakukan klasifikasi analisa sentimen. Beberapa penelitian yang peneliti jadikan sebagai referensi diantaranya yang dilakukan oleh Eka Rini Yulia *dkk*, bahwa algoritma *Naïve Bayes* merupakan metode yang cukup tepat untuk mengklasifikasi teks review pada *playstore* (Yulia & Solecha, 2021). Prancevicius *dkk*, menggunakan algoritma *Naïve Bayes*, *Random Forest*, *Decision Tree*, *Support Vector Machine*, dan *Logistics Regression* hasil eksperimen yang dilakukan *Naïve Bayes* mendapatkan akurasi 1-2% lebih besar dibanding algoritma yang lain [4]. Kemudian yang dilakukan Dedi D. Saputra *dkk*, bahwa *SMOTE* dapat melakukan optimasi terhadap algoritma *Naïve Bayes* sehingga mendapatkan akurasi 86.33% dibandingkan sebelumnya yang hanya sebesar 78.88%, hal ini karena *SMOTE* mampu menangani permasalahan ketidakseimbangan data [5]. Serta penelitian yang dilakukan oleh Latief *dkk* (2021) mendapatkan bahwa algoritma *Adaboost* dapat memberikan prediksi permasalahan *churn* yang lebih akurat dibandingkan algoritma *XGBoost* dan *random forest* [6]. Penelitian yang dilakukan Agus Byna *dkk* (2020), juga membuktikan bahwa terdapat peningkatan nilai akurasi algoritma *Naïve Bayes* setelah dilakukan optimasi dengan algoritma *Adaboost* [7]. Saifurrahman *dkk* juga melakukan penelitian dengan menganalisa sentimen dari aplikasi *duolingo* dengan metode *Naïve Bayes* dan *SMOTE*, terjadi peningkatan yang signifikan atas penggunaan *SMOTE* dalam penelitian, dampak kenaikan akurasi sebesar 14.81% [8].

Sentimen analisis yang sering juga dikenal *Opinion Mining* adalah suatu metode pendekatan analisis dalam hal mempelajari dan menganalisis tentang emosi, sikap, pandangan, dan pendapat seseorang dari sebuah teks. Pemikiran dasar dari analisis ini adalah dengan mendeteksi polaritas suatu teks dan melakukan klasifikasi. Polaritas ini dapat dikategorikan sebagai "*Positif*", "*Negatif*", atau "*Netral*" [9][10][11]. Salah satu kunci hal ini adalah *text mining* atau *Natural Language Processing* (NLP), seperti yang diketahui hal ini berkaitan dengan bagaimana

mengekstrak data atau makna dari sebuah teks [12]. Tujuan *text mining* adalah mencari informasi baru dari sebuah teks atau kalimat dan dalam melakukan ini diperlukan Analisa yang saling terhubung [13].

Naive Bayes Classifier adalah model sederhana dalam klasifikasi teks. Ini adalah model sederhana dari Bayesian network yang memiliki keunggulan diantaranya cepat, sederhana serta akurasi yang baik [14]. *Naive Bayes* memiliki ciri prediksi atau perkiraan yang kuat terhadap setiap kejadian atau kondisi [13]. *SMOTE* bertugas melakukan upsampling atas imbalanced data yang timbul dalam *dataset*. Sejak diciptakan pada tahun 2002, *SMOTE* telah terbukti berhasil diaplikasikan dengan beberapa domain yang berbeda. *SMOTE* juga telah menginspirasi beberapa pendekatan untuk mengatasi masalah ketidakseimbangan kelas [15]. *Adaboost* atau *Adaptive Boosting* salah satu boosting algoritma yang populer, langkah utamanya adalah membuat *weak learners* yang tidak memiliki korelasi dan menggabungkan prediksinya [6]. *Adaboost* tidak memberikan hasil yang maksimal secara individual, tetapi memberikan hasil yang maksimal secara bersama-sama. Penelitian ini menggunakan 2 aplikasi utama, yaitu *rapidminer* dan *Gataframework*. *Gataframework* adalah aplikasi *web* yang menyediakan layanan dalam hal *text processing* yang dapat membantu dalam proses *stemming* atau normalisasi kata dalam Bahasa Indonesia. Beberapa fitur yang digunakan dalam penelitian ini adalah annotation removal, transformation URL, tokenize regular expression, stemming Indonesian, dan Indonesian stop word removal [5][16].

2. Metode Penelitian

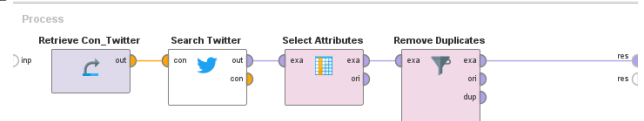


Gambar 1. Tahapan Penelitian

Penjelasan atas tahapan penelitian yang dilakukan diantaranya terdiri sebagai berikut:

Pengumpulan Data

Pengumpulan data dilakukan dengan metode Crawling data dari twitter dengan menggunakan tools yang ada pada aplikasi RapidMiner. Untuk mengakses data twitter, sebelumnya harus dilakukan koneksi ke API Twitter dengan cara login menggunakan akun Twitter yang dimiliki. Tools yang digunakan pada RapidMiner. Tools yang digunakan dapat dilihat pada gambar berikut :



Gambar 2. Crawling Twitter Tools Rapidminer

Preprocessing Tahap 1 (Gataframework)

Fitur yang digunakan pada aplikasi *Gataframework* dapat dilihat pada gambar berikut:



Gambar 3. Fitur *Gataframework*

Penjelasan atas fitur yang digunakan adalah sebagai berikut:

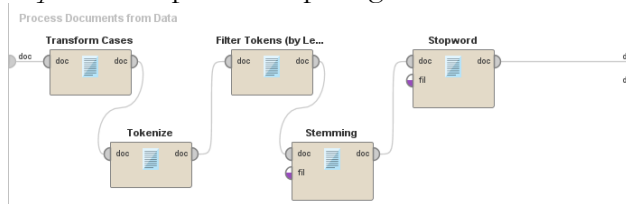
- Annotation Removal**
Tahapan ini menghilangkan anotasi yang ada dalam *dataset*.
- Transformation: Remove URL**
Setelah anotasi dihilangkan, selanjutnya dilakukan penghapusan terhadap link atau URL yang ada pada *dataset*.
- Tokenize Regular Expression (Regexp)**
Proses ini dilakukan untuk menghilangkan tanda baca atau simbol atau karakter khusus lain
- Indonesian Stemming**
Proses stemming adalah mengubah kata yang ada pada *dataset* yang memiliki imbuhan, menjadi kata dasar berdasarkan bahasa Indonesia menggunakan database kata dasar yang ada pada aplikasi *Gataframework*.

e. Indonesian Stop Word Removal

Proses ini dilakukan untuk menghilangkan kata-kata yang tidak relevan untuk digunakan untuk penilaian sentimen seperti kata untuk, dengan, lalu, karena kata-kata ini tidak memiliki makna apabila berdiri sendiri. Kata tersebut perlu kata lain sehingga memiliki makna. Kamus data kata stop word menggunakan basis data yang ada pada aplikasi *Gataframework*.

Preprocessing Tahap 2 (Rapidminer)

Pemodelan preprocessing tahap 2 pada aplikasi *Rapidminer* dapat dilihat pada gambar berikut:



Gambar 4. Pemodelan *Preprocessing* tahap 2 (*Rapidminer*)

Alat yang digunakan dalam penelitian pada aplikasi *Rapidminer* dalam melakukan preprocessing tahap 2 ini diantaranya,

a. Transform Case

Operator ini digunakan untuk merubah semua karakter atau huruf dalam *dataset* menjadi sama (misal: menjadi huruf besar semua atau huruf kecil semua).

b. Tokenization

Operator ini digunakan untuk memecah satu rangkaian teks kalimat menjadi bentuk per kata. Istilah atas kata yang sudah dipecah biasa disebut token.

c. Filter Token (By Length)

Operator ini digunakan untuk melakukan filter kata yang ada sesuai parameter yang akan kita tentukan. Pada penelitian ini parameter yang digunakan adalah kata yang terdiri minimal 3 karakter dan maksimal 25 karakter.

d. Stemming

Stemming ini digunakan untuk lebih melengkapi hasil stemming yang sudah dilakukan pada *Gataframework*.

e. Stopwords

Operator ini digunakan untuk menghilangkan kata yang tidak diperlukan atau tidak memiliki arti yang tidak mempengaruhi hasil penelitian. Ini

digunakan untuk mengurangi waktu proses dalam *Cross Validation* sehingga menjadi lebih efektif.

Pengujian dengan Algoritma Naive Bayes, SMOTE dan Adaboost

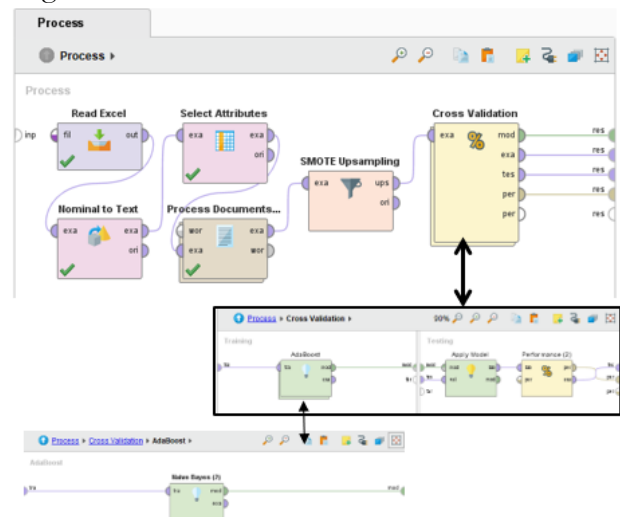
Tahap ini dilakukan pengujian kombinasi algoritma yang dapat menghasilkan kualifikasi terbaik yang dapat digunakan. Ujicoba kombinasi algoritma digunakan untuk mendapatkan data kualifikasi yang dapat digunakan dalam proses analisa sentimen.

Evaluation

Pada tahap ini performa kombinasi algoritma diukur berdasarkan sistem. Hasil evaluasi ini akan ditampilkan dalam bentuk *confussion matrix* yang terdiri dari *accuracy*, *precision*, dan *recall* serta kualitas hasil modelling dapat dilihat lewat *AUC*.

3. Hasil dan Pembahasan

Hasil penelitian berdasarkan metode penelitian yang telah dijelaskan sebelumnya diantaranya adalah sebagai berikut:



Gambar 5. Proses dalam *rapidminer*

Pengumpulan Data

Tweet yang dikumpulkan adalah tweet yang ada pada akun *twitter* @bankbtn. Populasi dalam penelitian ini adalah seluruh komentar/*tweet* yang ada pada akun @bankbtn, dan sampel yang digunakan adalah sebanyak 500 data tweet yang diambil dalam rentang waktu Desember 2021 – Maret 2022. Selanjutnya tweet tersebut dilakukan pelabelan dengan sentiment

“Positive” dan “Negative”. Hasil crawling dan pelabelan dapat dilihat pada gambar berikut:

N	Text	Status
1	"Solusi yg dtawarkn ini bgmn kita mensinergikan tanah yg disediakan o/ @KAI121,Perumnas yg mbangun dbntu oleh BUMN Karya,@bankbtn yg mberikn financing dbntu o/ Jasa Keuangan,@pln_123& @TelkomIndonesia bsineri mberikn solusi yg tbaik, -@erickthohir- #Eri	positive
3	@agnesriani @bankbtn_syariah @bankbtn Hal Sahabat BTN Syariah @agnesriani mengenai hal tersebut kami sarankan untuk kakak dalam melakukan pengininan data silahkan dilakukan di Kantor Cabang BTN Syariah terdekat dengan membawa Buku Tabungan, Kartu ATM, dan KTP yang masih berlaku. Tks*Sasa	positive
4	@AngkuKayo Ngadi2 emang @bankbtn ini. Untah darat banget	negative
5	@aninditop12 @bankbtn Siap kak	positive

Gambar 6. Dataset hasil *Crawling* dan *Labelling*

Preprocessing Tahap 1 (Gataframework)

a. Annotation Removal

Tabel 1. Hasil Annotation Removal

Text	@Anotation Removal
"Solusi yg dtawarkn ini bgmn kita mensinergikan tanah yg disediakan o/ @KAI121,Perumnas yg mbangun dbntu oleh BUMN Karya,@bankbtn yg mberikn financing dbntu o/ Jasa Keuangan,@pln_123& @TelkomIndonesia bsineri mberikn solusi yg tbaik, - @erickthohir- #Eri	"solusi yg dtawarkn ini bgmn kita mensinergikan tanah yg disediakan o/ yg mbangun dbntu oleh bumn karya, yg mberikn financing dbntu o/ jasa keuangan, bsineri mberikn solusi yg tbaik, - #eri

b. Transformation: Remove URL

Tabel 2. Hasil Transformation: Remove URL

@Anotation Removal	Transformation: Remove URL
"solusi yg dtawarkn ini bgmn kita mensinergikan tanah yg disediakan o/ yg mbangun dbntu oleh bumn karya, yg mberikn financing dbntu o/ jasa keuangan, bsineri mberikn solusi yg tbaik, - #eri	"solusi yg dtawarkn ini bgmn kita mensinergikan tanah yg disediakan o/ yg mbangun dbntu oleh bumn karya, yg mberikn financing dbntu o/ jasa keuangan, bsineri mberikn solusi yg tbaik, - #eri

c. Tokenize Regular Expression (Regexp)

Tabel 3. Hasil Regexp

Transformation: Remove URL	Regexp
"solusi yg dtawarkn ini bgmn kita mensinergikan tanah yg disediakan o/ yg mbangun dbntu oleh bumn karya, yg mberikn financing dbntu o/ jasa keuangan, bsineri mberikn solusi yg tbaik, - #eri	solusi yg dtawarkn ini bgmn kita mensinergikan tanah yg disediakan o yg mbangun dbntu oleh bumn karya yg mberikn financing dbntu o jasa keuangan bsineri mberikn solusi yg tbaik eri

d. Indonesian Stemming

Tabel 4. Hasil Indonesian Stemming

Regexp	Indonesian Stemming
solusi yg dtawarkn i ni bgmn kita mensinergikan tanah yg disediakan o yg mbangun dbntu oleh bumn karya yg mberikn financing dbntu o jasa keuangan bsineri mberikn solusi yg tbaik eri	solusi yg dtawarkn ini bgmn kita sinergi tanah yg sedia o yg mbangun dbntu oleh bumn karya yg mberikn financing dbntu o jasa uang bsineri mberikn solusi yg tbaik eri

e. Indonesian Stop Word Removal

Tabel 5. Hasil Indonesian Stop Word Removal

Indonesian Stemming	Indonesian Stop word removal
solusi yg dtawarkn ini bgmn kita sinergi tanah yg sedia o yg mbangun dbntu oleh bumn karya yg mberikn financing dbntu o jasa uang bsineri mberikn solusi yg tbaik eri	solusi dtawarkn bgmn sinergi tanah sedia mbangun dbntu bumn karya mberikn financing dbntu jasa uang bsineri mberikn solusi tbaik eri

*Preprocessing Tahap 2 (Rapidminer)**a. Transform Case*

Tabel 6. Hasil Transform Case

Indonesian <i>Stop Word Removal</i>	<i>Transform Case</i>
solusi dtawarkn bgmn sinergi tanah sedia mbangun dbntu bum karya mberikn financing dbntu jasa uang bsinerigi mberikn solusi tbaik eri	solusi dtawarkn bgmn sinergi tanah sedia mbangun dbntu bum karya mberikn financing dbntu jasa uang bsinerigi mberikn solusi tbaik eri

b. Tokenize

Tabel 7. Hasil Tokenize

<i>Transform Case</i>	<i>Tokenize</i>
solusi dtawarkn bgmn sinergi tanah sedia mbangun dbntu bum karya mberikn financing dbntu jasa uang bsinerigi mberikn solusi tbaik eri	solusi dtawarkn bgmn sinergi tanah sedia mbangun dbntu bum karya mberikn financing dbntu jasa uang bsinerigi mberikn solusi tbaik eri

c. Filter Tokens (by Length)

Tabel 8. Hasil Filter Tokens (By Length)

<i>Tokenize</i>	<i>Filter Token (by Length)</i>
solusi dtawarkn bgmn sinergi tanah sedia mbangun dbntu bum karya	solusi dtawarkn bgmn sinergi tanah sedia mbangun dbntu bum karya

mberikn financing dbntu jasa uang bsinerigi mberikn solusi tbaik eri	mberikn financing dbntu jasa uang bsinerigi mberikn solusi tbaik eri
---	---

d. Stemming

Tabel 9. Hasil Stemming tambahan

<i>Filter Token (by Length)</i>	<i>Stemming</i>
solusi dtawarkn bgmn sinergi tanah sedia mbangun dbntu bum karya mberikn financing dbntu jasa uang bsinerigi mberikn solusi tbaik eri	solusi tawar bagaimana sinergi tanah sedia bangun bantu bum karya beri financing bantu jasa uang bersinerigi beri solusi terbaik eri

e. Stopwords

Tabel 10. Hasil Stopwords tambahan

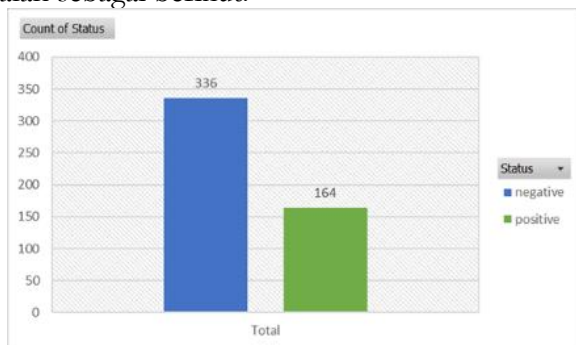
<i>Stemming</i>	<i>Stopwords</i>
solusi dtawarkn bgmn sinergi tanah sedia mbangun dbntu bum karya mberikn financing	solusi tawar bagaimana sinergi tanah sedia bangun bantu bum karya beri financing

dbntu	bantu
jasa	jasa
uang	uang
bsinergi	bersinergi
mberikn	beri
solusi	solusi
tbaik	terbaik
eri	eri

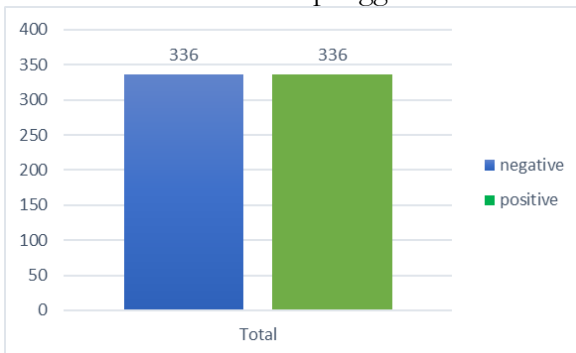
<i>Naïve Bayes</i>	11	70	53	266
	1			
<i>SMOTE + Naïve Bayes</i>	31	70	25	266
	1			
<i>SMOTE + Adboost + Naïve Bayes</i>	30	57	30	279
	6			

Penggunaan SMOTE

Penggunaan *SMOTE* dapat menangani *dataset* yang statusnya imbalance sehingga dapat memberikan hasil klasifikasi yang lebih baik. Perbedaan *dataset* sebelum dan setelah diberikan operator *SMOTE* adalah sebagai berikut:



Gambar 7. *Dataset* sebelum penggunaan *SMOTE*



Gambar 8. *Dataset* setelah penggunaan *SMOTE*

Evaluation

Setelah dilakukan tahapan sebelumnya, selanjutnya dilakukan perhitungan hasil kualifikasi dan percobaan kombinasi algoritma untuk mendapatkan hasil kualifikasi yang paling baik. Hasil *evaluation* dapat dilihat pada tabel berikut:

Tabel 11. Hasil Confussion Matrix

Algoritma	T P	F P	F N	T N
Adboost + <i>Naïve Bayes</i>	10 7	54	57	282

Tabel 12. Hasil Evaluasi

Algoritma	Accu racy	Precis ion	Rec all	A U C
Adboost + <i>Naïve Bayes</i>	77.80 %	83.27 %	83.9 1%	0.7 59
<i>Naïve Bayes</i>	75.40 %	83.49 %	79.1 4%	0.6 14
<i>SMOTE + Naïve Bayes</i>	85.87 %	91.66 %	79.1 4%	0.8 63
<i>SMOTE + Adboost + Naïve Bayes</i>	87.05 %	90.63 %	83.0 0%	0.9 09

4. Kesimpulan dan Saran

Kondisi imbalance *dataset* cukup berpengaruh pada hasil kualifikasi algoritma, terbukti apabila tanpa ditambahkan operator *SMOTE*, hasil kualifikasi yang didapat masih dibawah 78% akurasi. Hasil penelitian juga membuktikan operator *SMOTE* memang efektif untuk mengatasi kondisi ini, terbukti dengan kenaikan *accuracy* menjadi sebesar 85.87%. Hasil penelitian juga didapat, kombinasi algoritma *SMOTE*, *Naïve Bayes* dan *Adaboost* terbukti memberikan hasil kualifikasi terbaik, dengan akurasi sebesar 87.05%, *precision* 90.63%, *recall* 83.00% serta AUC 0.909. Hasil ini sudah cukup apabila diterapkan menjadi sebuah mesin klasifikasi. Saran dalam penelitian lanjutan adalah memperbanyak jumlah *dataset* yang ada dan melakukan pengujian terhadap algoritma yang telah diuji untuk melihat kehandalan kombinasi algoritma ini serta dilakukan penerapan menjadi sebuah mesin klasifikasi.

5. Daftar Pustaka

- [1] Augustinah, F. & W., 2019. Pemanfaatan Media Sosial Sebagai Sarana Promosi Makanan Ringan Kripiki Singkong Di Kabupaten Sampang. *Jurnal Dialektika*, 4(2), pp. 1-20.

- [2] Marpaung, J., 2018. Pengaruh Penggunaan Gadget Dalam Kehidupan. *Jurnal KOPASTA*, 5(2), pp. 55-64.
- [3] Kemp, S., 2022. *Digital 2022: Indonesia*. [Online] Available at: <https://datareportal.com/reports/digital-2022-indonesia> [Accessed 08 April 2022].
- [4] Pranckevicius, T. & Marcinkevicius, V., 2017. Comparison of Naïve Bayes, Random Forest, Decision Tree, Support Vector Machines, and Logistic Regression Classifiers for Text Reviews Classification. *Baltic J. Modern Computing*, 5(2), pp. 221-232.
- [5] Saputra, D. D. et al., 2019. Optimization Sentiments of Analysis from Tweets in myXLCare using Naive Bayes Algorithm and Synthetic Minority Oversampling Technique Method. *Journal Of Physics: Conference Series*, Volume 1471, pp. 1-7.
- [6] Latief, I. M., Subekti, A. & Gata, W., 2021. PREDIKSI TINGKAT PELANGGAN CHURN PADA PERUSAHAAN TELEKOMUNIKASI DENGAN ALGOITMA ADABOOST. *Jurnal Informatika*, 21(1), pp. 34-43.
- [7] Byna, A. & Basit, M., 2020. Penerapan Metode Adaboost Untuk Mengoptimasi Prediksi Penyakit Stroke Dengan Algoritma Naïve Bayes. *Jurnal SISFOKOM (Sistem Informasi dan Komputer)*, 09(03), pp. 407-411.
- [8] Chohan, S., Nugroho, A., Bayu Aji, A. M. & Gata, W., 2020. Analisis Sentimen Aplikasi Duolingo Menggunakan Metode Naïve Bayes dan Synthetic Minority Over Sampling Technique. *Paradigma – Jurnal Informatika dan Komputer*, 22(2), pp. 139-144.
- [9] Wongkar, M. & Angdresey, A., 2019. Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler : Twitter. *2019 Fourth International Conference on Informatics and Computing (ICIC)*, pp. 1-5.
- [10] Kanimozhi, P. & Elavarasi, D., 2018. A Survey on Sentiment Analysis using Twitter Dataset. *International Journal of Engineering Research & Technology (IJERT)*, 6(14), pp. 1-3.
- [11] Alsaeedi, A. & Khan, M. Z., 2019. A Study on Sentiment Analysis Techniques of Twitter Data. *(IJACSA) International Journal of Advanced Computer Science and Applications*, 10(2), pp. 1-2.
- [12] Swamynathan, M., 2017. *Mastering Machine Learning with Python in Six Steps*. New York: Apress Media, LLC.
- [13] Ramadhani, S.H. and Wahyudin, M.I., 2022. Analisis Sentimen Terhadap Vaksinasi Astra Zeneca pada Twitter Menggunakan Metode Naïve Bayes dan K-NN. *Jurnal JTik (Jurnal Teknologi Informasi dan Komunikasi)*, 6(4), pp.526-534.
- [14] Indrayuni, E., 2019. Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes. *JURNAL KHATULISTIWA INFORMATIKA*, VII(1), pp. 29-36.
- [15] Fernandez, A., Garcia, S., Herrera, F. & Chawla, N. V., 2018. SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary. *Journal of Artificial Intelligence Research*, Issue 61, pp. 863-905.
- [16] Yulia, E. R. & Solecha, K., 2021. Implementasi Particel Swarm Optimization (PSO) pada Analysis Sentiment Review Aplikasi Trafi menggunakan Algoritma Naive Bayes (NB). *Jurnal Teknik Komputer AMIK BSI*, 7(1), pp. 25-29.