

## Optimasi Analisis Sentimen Pada *Twitter* Olshop Tokopedia Menggunakan Textmining Dengan Algoritma *Naïve Bayes & Adaboost*

Hartati<sup>1</sup>, Deni Hermawan<sup>2</sup>, M. Akhsanal<sup>3</sup>, Zailani Wahyudi<sup>4</sup>, Angga Ariyanto<sup>5</sup>, Dedi Dwi Saputra<sup>6</sup>

<sup>1,5</sup>Mahasiswa Sistem Informasi, Fakultas Teknologi Informasi, Universitas Nusa Mandiri, Indonesia

<sup>6</sup>Dosen DWBI, Fakultas Teknologi Informasi, Universitas Nusa Mandiri, Indonesia

e-mail: hartati2890@gmail.com<sup>1</sup>, denihermawan173@gmail.com<sup>2</sup>, makhsanal@gmail.com<sup>3</sup>, zailaniwahyudi7@gmail.com<sup>4</sup>, ariyanto.angga52@gmail.com<sup>5</sup>, dedi.eis@nusamandiri.ac.id<sup>6</sup>

### Abstract

*Sentiment Analysis or commonly called Opinion Mining is the process of understanding, extracting and processing textual data automatically to obtain sentiment information contained in a sentence of opinion or opinion on a problem or object by someone, whether it tends to have a positive or negative opinion. This study aims to classify tweet data into 2 classifications, namely positive and negative. In this study, Indonesian text is used on Twitter social media in the form of tweets related to Tokopedia. Public opinion contained in the tweet can be used as material to find out whether tweets on Twitter, especially on Tokopedia, are classified as positive or negative. The data used consists of 1,000 tweet data. This dataset comes from the tweets of Tokopedia customers written on the Tokopedia twitter account. In text mining techniques, "transform case", "tokenize", "token filter by length", "stemming" are used to build classifications. Gataframework is used to help during the preprocessing and cleansing process. RapidMiner is used to help create sentiment analysis in comparing three different classification methods, on Tokopedia's tweet data. The method used to compare in this research is the Naïve Bayes algorithm and the Naïve Bayes algorithm which is added with the Synthetic Minority Over-sampling Technique (SMOTE) feature and the Naïve Bayes algorithm is added with the Synthetic Minority Over-sampling Technique (SMOTE) feature which is optimized with Adboost. . The Naïve Bayes algorithm added with the Synthetic Minority Over-sampling Technique (SMOTE) feature, which was optimized with Adboost, got the best score. With 94.95% accuracy, 90.86% precision, 100.00% recall and 0.950 AUC.*

**Keywords:** Text Mining, Adaboost, Naive Bayes, Classification, Sentiment Analysis

### Abstrak

*Sentimen Analisis atau biasa disebut Opinion Mining merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk memperoleh informasi sentimen yang terkandung dalam suatu kalimat opini atau pendapat terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung beropini positif atau negatif. Penelitian ini bertujuan untuk mengklasifikasi data tweet menjadi 2 klasifikasi yaitu positif dan negatif. Dalam penelitian ini menggunakan teks Bahasa Indonesia yang terdapat di sosial media twitter berupa tweet yang berkaitan dengan tokopedia. Opini masyarakat yang terdapat di tweet tersebut dapat dimanfaatkan sebagai bahan untuk mengetahui apakah tweet pada twitter khususnya pada tokopedia masuk kedalam klasifikasi positif atau negatif. Data yang digunakan terdiri dari 1.000 data tweet. Dataset ini berasal dari tweet pelanggan tokopedia yang dtulis diakun twitter Tokopedia. Dalam teknik text mining, "transform case", "tokenize", "token filter by length", "stemming" digunakan untuk membangun klasifikasi. Gataframework digunakan untuk membantu selama proses preprocessing dan cleansing. RapidMiner digunakan untuk membantu menciptakan sentimen analisis dalam membandingkan tiga metode klasifikasi yang berbeda, pada data tweet tokopedia. Metode yang digunakan untuk membandingkan pada penelitian ini adalah, algoritma Naïve Bayes dan algoritma Naïve Bayes yang ditambahkan feature Synthetic Minority Over-sampling Technique*



(SMOTE) serta algoritma Naive Bayes yang ditambahkan feature Synthetic Minority Over-sampling Technique (SMOTE) yang di optimasi dengan Adaboost. Algoritma Naive Bayes yang ditambahkan feature Synthetic Minority Over-sampling Technique (SMOTE), yang dioptimasi dengan Adaboost mendapatkan nilai terbaik. Dengan nilai akurasi 94.95%, Presisi 90.86%, Recall 100.00% dan AUC 0,950.

**Kata Kunci:** Text Mining, Adaboost, Naive Bayes, Klasifikasi, Sentimen Analisis

## 1. PENDAHULUAN

Belakangan ini, situs jejaring sosial seperti *twitter*, facebook, Instagram, blog banyak digunakan masyarakat untuk mengekspresikan sentiment, pandangan masukan, pendapat, sarana promosi, iklan, kampanye dan lain-lain [1]. Pada *twitter* Tokopedia terdapat isu-isu yang terkait dengan pelayanan ataupun Tokopedia itu sendiri yang memunculkan banyak opini masyarakat, tidak hanya opini yang positif tetapi juga opini yang bersifat negative [2]. Oleh karena itu, *Twitter* dianggap lebih diminati oleh masyarakat Indonesia karena dirasa lebih mudah dan simpel dalam mempresentasikan opini public [2].

Text Mining adalah proses ekstraksi pola berupa informasi dan pengetahuan yang berguna dari sejumlah besar sumber data teks, seperti dokumen Word, PDF, kutipan teks, dll. Jenis masukan untuk penambangan teks ini disebut data tak terstruktur dan merupakan pembeda utama dengan penambangan data yang menggunakan data terstruktur atau basis data sebagai masukan. Penambangan teks dapat dianggap sebagai proses dua tahap yang diawali dengan penerapan struktur terhadap sumber data teks dan dilanjutkan dengan ekstraksi informasi dan pengetahuan yang relevan dari data teks terstruktur ini dengan menggunakan teknik dan alat yang sama dengan penambangan data. Proses yang umum dilakukan oleh penambangan teks di antaranya adalah perangkuman otomatis, kategorisasi dokumen, penggugusan teks, deteksi plagiarisme, dll [3]. Klasifikasi data adalah proses mengklasifikasikan data ke dalam beberapa kategori untuk penggunaan yang paling efektif dan produktif. Tujuan klasifikasi adalah untuk memprediksi kelas target secara akurat untuk setiap kasus data. Sebuah algoritma yang khusus digunakan untuk mengimplementasikan klasifikasi dikenal dengan classifier. Istilah “pengklasifikasian” terkadang juga mengacu pada fungsi matematika, yang diimplementasikan oleh algoritma klasifikasi [1].

Analisis sentimen (juga dikenal sebagai penggalian opini atau AI emosi) adalah penggunaan pemrosesan bahasa alami, analisis teks, linguistik komputasi, dan biometrik untuk mengidentifikasi, mengekstrak, mengukur, dan mempelajari keadaan afektif dan informasi subjektif secara sistematis. Analisis sentimen diterapkan secara luas pada materi suara pelanggan seperti ulasan dan tanggapan survei, media online dan sosial, dan materi perawatan kesehatan untuk aplikasi yang berkisar dari pemasaran hingga layanan pelanggan hingga pengobatan klinis. Dengan munculnya model bahasa yang dalam, seperti RoBERTa, domain data yang lebih sulit juga dapat

dianalisis, misalnya, teks berita di mana penulis biasanya mengungkapkan pendapat/sentimen mereka secara kurang eksplisit [4].

## 2. METODOLOGI PENELITIAN

Analisis Sentimen merupakan solusi dari masalah yang telah dipaparkan sebelumnya. Pada penelitian ini algoritma yang akan digunakan adalah *Naïve Bayes Classifier* (NBC) & Adaboost. Algoritma *Naïve Bayes* dan Adaboost dirasa cocok untuk menganalisis sentiment dikarenakan algoritma ini bertujuan sebagai metode klasifikasi dalam kategori positif dan negatif.

Bagian ini berisi tinjauan penelitian yang sebelumnya dilakukan di bidang analisis sentimen untuk data langsung. Banyak penelitian telah dilakukan sampai saat ini dalam bidang ini. Data dari pengguna di media sosial untuk mengekstrak sentimen orang terhadap topik, produk, tren, dll. Studi ini berfokus pada penggalian informasi yang berguna dari bahasa alami pengguna dan memprosesnya untuk mendapatkan sentimen nyata dari bahasa tersebut.

Berdasarkan latar belakang yang dikemukakan di atas maka dapat dirumuskan masalah penelitian sebagai berikut: Bagaimana implementasi Algoritma *Naive Bayes Classifier* (NBC) & Adaboost dalam mengklasifikasikan data tweet pada media sosial *twitter* apakah data tweet pada tokopedia tersebut termasuk ke dalam kelas bersentimen positif atau negatif ?

Selain itu agar masalah yang sedang di tinjau lebih terarah dan tercapai sasaran yang telah ditentukan, maka harus diberikan batasan-batasan masalah antara lain:

- a) Objek yang digunakan pada penelitian ini adalah data tweet dari pengguna *twitter* Tokopedia April 2022.
- b) Tweet yang akan digunakan yaitu kalimat tweet yang menggunakan bahasa Indonesia saja.
- c) Algoritma yang akan di gunakan untuk pengklasifikasian dalam penelitian ini adalah Adaboost dan NBC (*Naive Bayes Classifier*).
- d) Pada penelitian ini proses stemming dan stopword hanya untuk kata-kata berbahasa indonesia saja.

Adapun tujuan pada penelitian ini adalah untuk mengklasifikasi sentiment pada data tweet dan menggunakan text mining dengan metode komparatif algoritma *Naïve Bayes* & Adaboost dan algoritma untuk analisis sentiment di media sosial *twitter*. Adapun manfaat pada penelitian ini adalah sebagai berikut :

- a) Penelitian ini diharapkan dapat memberikan tambahan pengetahuan tentang implementasi algoritma *Naïve Bayes* (NBC) & Adaboost dapat memberikan gambaran mengenai bagaimana analisis sentimen pada jejaring sosial *twitter* untuk mengklasifikasi opini untuk berbagai kepentingan serta pengoptimalan informasi jejaring sosial untuk kepentingan publik.

- b) Selain itu juga bermanfaat untuk mencari sebuah informasi tentang suatu opini atau pendapat mempunyai respon positif atau negatif dari publik khususnya pengguna *twitter*.
- c) Penelitian ini juga bermanfaat untuk mengetahui berapa keberhasilan algoritma algoritma *Naïve Bayes* (NBC) & Adaboost pada pengklasifikasian teks berbahasa Indonesia.

Penelitian sebelumnya yang pernah dilakukan berkaitan dengan masalah yang dihadapi penulis lainnya yaitu dengan judul “ Analisis Sentiment *Twitter* Menggunakan Text Mining dengan Algoritma *Naïve Bayes Classifier* “ oleh Adhi Viky Sudiantoro dan Eri Zuliarso pada tahun 2018. Pada Penelitian tersebut hasil pengujian dan analisis yang dilakukan oleh penulis dengan metode super vector machine dengan akurasi mencapai 84 %, hasil tersebut dipengaruhi oleh beberapa hal seperti komposisi jumlah data training dn testing, jumlah dataset yang digunakan dan komposisi jumlah data positif dan negatif [2].

Penelitian lainnya adalah “ Optimization Sentiments of Analysis from Tweets in myXLCare using *Naïve Bayes* Algorithm and Synthetic Minority over Sampling Technique Method “ oleh Dedi Dwi Saputra yang ditulis kedalam bahasa inggris pada tahun 2020, dengan hasil akurasi 86.33%, presisi 82.58 %, serta recall 92.38% [5].

Peneletian yang serupa lainnya adalah “Sentiment Analysis and Classification of Tweets Using Data Mining “ yang ditulis oleh Md Shoeb dan Jawed Ahmed dalam bahasa inggris pada tahun 2017 dengan hasil akurasi 64.42 % [6].

Penelitian lainnya adalah “Analisi Sentiment Aplikasi Ruang Guru di *Twitter* menggunakan Algoritma Klasifikasi “ yang ditulis oleh Angelina Puput Giovani, Ardiansyah, Tuti Haryanti, Laela Kurniawati, Windu Gata Pada Tahun 2020 dengan sebesar 78,55% dan AUC sebesar 0,853 [7].

### 2.1. Analisis Data

Pada penelitian ini, data yang digunakan ialah tweet berbahasa Indonesia yang terdapat pada *Twitter*. Tweet yang digunakan yaitu tweet-tweet yang mengandung opini/pendapat masyarakat Indonesia terhadap tokopedia. Tweet yang digunakan sebagai data berjumlah 1.000. kemudian data dipilih secara manual yaitu dengan cara memilih kalimat-kalimat tweet yang berbahasa Indonesia dan tidak mengandung gambar. Data yang terpilih kemudian disimpan dalam bentuk excel. Data dalam penelitian ini terdiri dari dua jenis yaitu data latih dan data uji. Untuk keperluan data latih , data yang telah dikumpulkan dikategorikan secara manual dan menilai sentiment dalam tweet tersebut masuk dalam kategori positif atau negatif. Data yang dipakai untuk data latih yaitu sebanyak 1.000 data tweet, dimana data itu terbagi menjadi 2 bagian yaitu data positif dan data negatif.

**Tabel 1.** Tabel Jenis Sentimen

Ket	Sentimen		
	Negatif	Positif	Grand Total
Total	59	941	1.000

Dari tabel 1 data yang mengandung kata positif 941 data dan kata negatif 59 data.

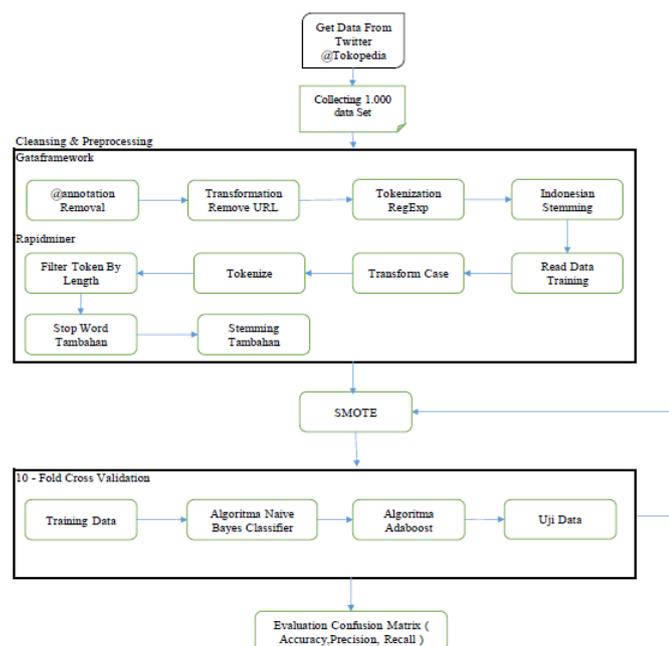
## 2.2. Analisis Pemrosesan Teks dan Implementasi Algoritma *Naïve Bayes*

Pemrosesan teks merupakan proses menggali, mengolah, mengatur informasi dengan cara menganalisis hubungannya, aturan-aturan yang ada di data tekstual semi terstruktur atau tidak terstruktur. Untuk lebih efektif dalam proses pemrosesan teks dilakukan langkah transformasi data ke dalam suatu format yang memudahkan untuk kebutuhan peneliti. Proses ini disebut text processing. Setelah dalam bentuk yang lebih terstruktur dengan adanya proses diatas data dapat dijadikan sumber data yang dapat diolah lebih lanjut. Tahapan untuk Pemrosesan teks terdiri dari tokenizing, normalisasi fitur, case folding dan stopword removal.

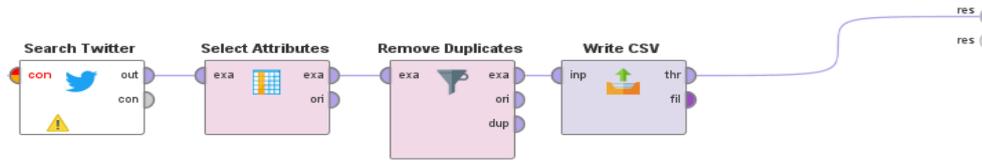
Data yang sudah melalui proses pemrosesan teks data kemudian akan melalui tahap klasifikasi menggunakan *Naïve Bayes Classifier* untuk mengetahui data tersebut apakah berkategori positif atau berkategori negatif. Metode *Naïve Bayes Classifier* digunakan untuk pengkategorian, yaitu untuk mendeteksi apakah pendapat/opini terhadap sebuah masalah atau objek oleh seseorang cenderung berkategori negatif atau positif.

## 2.3. Metode Penelitian

Metode penelitian yang akan digunakan adalah pengumpulan data *tweet*. Data *Tweet* diambil dengan metode *Crawling* dari *Twitter*. Data yang diambil hanya *tweet* dalam bahasa Indonesia yaitu 1.000 *tweet* dengan kata-kata tokopedia. Data diambil secara acak baik dari akun pengguna biasa maupun akun media online di *Twitter*.



**Gambar 1.** Metode Penelitian untuk Analisis Sentimen



**Gambar 2.** Getting Data dari *Twitter*

### 3. HASIL DAN PEMBAHASAN

#### 3.1. Pembahasan

Menyiapkan data yang akan diolah di Rapidminer, dalam penelitian ini data diperoleh dari getting data *twitter* yang berupa data teks tweet yang berhubungan dengan tokopedia pada April 2022. Data tersebut di simpan dalam bentuk excel, kemudian data tersebut di upload ke dalam aplikasi Rapid Miner untuk diolah lebih lanjut. Berikut sebagian data tweet yang berhubungan dengan tokopedia:

No	Text	Sentiment
1	JAY dan JUNGWON mau ingetin kamu untuk nonton #TokopediaTVShow Ramadan Ekstra 5 hari lagi!! Siapa yang udah ngga sabar nonton ENHYEPEN? Catat jadwalnya : ? Rabu, 20 April 2022 ? 19:30-21:00 WIB ? NET TV, Indosiar, Trans 7, ANTV, Tokopedia PLAY	
2	ENHYEPEN #TokopediaENHYEPEN Sayang ENGENE banyak-banyak?? Sayang ENGENE banyak-banyak?? Sayang ENGENE banyak-banyak?? Sayang ENGENE banyak-banyak?? Sayang ENGENE banyak-banyak??	positif
3	#TokopediaENHYEPEN <a href="https://t.co/NBSggzT7Tj">https://t.co/NBSggzT7Tj</a> Besok lagi siapa yahh ??	positif
4	#TokopediaENHYEPEN <a href="https://t.co/3Fn1srFXSN">https://t.co/3Fn1srFXSN</a>	positif
5	Coba cek ini, deh: 'Avoskin Miraculous Retinol Ampoule 30ml' Rp184.200 di toko Avoskin Tokopedia. Bisa Bebas Ongkir. Iho. <a href="https://t.co/bRbcT0XyU">https://t.co/bRbcT0XyU</a>	positif
6	Coba cek ini, deh: 'FOCALLURE Jelly-Clear Dewy Lip Tint-Lip Gloss Lipstick FA208 - #203' Rp25.000 di toko Focallure Official Store Tokopedia. Bisa Bebas Ongkir. Iho. <a href="https://t.co/lmC5VbkrG">https://t.co/lmC5VbkrG</a>	positif
7	#Tokopedia: Ding Dong Deng! 5 hari lagi jangan lupa nonton yaah ENGENE ????	
8	#TokopediaENHYEPEN <a href="https://t.co/FeUGGpsWo">https://t.co/FeUGGpsWo</a>	positif

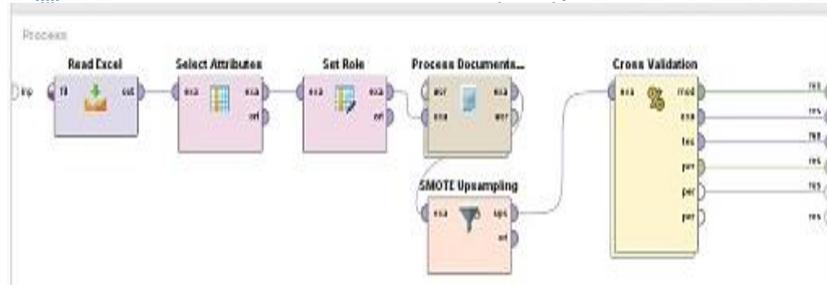
**Gambar 3.** Data *Tweet*

Selanjutnya Cleansing 1.000 data tweet yang sudah melalui tahap pemrosesan teks dengan *gataframework*.

A	B	C	D	E	F	G	H
No	Text	Status	@(notation Removal)	Transformation: Remove URL	Regex	Indonesian Stemming	Indonesian Stop word removal
1	JAY dan JUNGWON mau ingetin kamu untuk nonton #TokopediaTVShow Ramadan Ekstra 5 hari lagi!! Siapa yang udah ngga sabar nonton ENHYEPEN? Catat jadwalnya : ? Rabu, 20 April 2022 ? 19:30-21:00 WIB ? NET TV, Indosiar, Trans 7, ANTV, Tokopedia PLAY	positif	@(mention Removal)	Transformation: Remove URL	Regex	Indonesian Stemming	Indonesian Stop word removal
2	ENHYEPEN #TokopediaENHYEPEN Sayang ENGENE banyak-banyak?? Sayang ENGENE banyak-banyak?? Sayang ENGENE banyak-banyak?? Sayang ENGENE banyak-banyak??	positif	@(mention Removal)	Transformation: Remove URL	Regex	Indonesian Stemming	Indonesian Stop word removal
3	#TokopediaENHYEPEN <a href="https://t.co/NBSggzT7Tj">https://t.co/NBSggzT7Tj</a> Besok lagi siapa yahh ??	positif	@(mention Removal)	Transformation: Remove URL	Regex	Indonesian Stemming	Indonesian Stop word removal
4	#TokopediaENHYEPEN <a href="https://t.co/3Fn1srFXSN">https://t.co/3Fn1srFXSN</a>	positif	@(mention Removal)	Transformation: Remove URL	Regex	Indonesian Stemming	Indonesian Stop word removal
5	Coba cek ini, deh: 'Avoskin Miraculous Retinol Ampoule 30ml' Rp184.200 di toko Avoskin Tokopedia. Bisa Bebas Ongkir. Iho. <a href="https://t.co/bRbcT0XyU">https://t.co/bRbcT0XyU</a>	positif	@(mention Removal)	Transformation: Remove URL	Regex	Indonesian Stemming	Indonesian Stop word removal
6	Coba cek ini, deh: 'FOCALLURE Jelly-Clear Dewy Lip Tint-Lip Gloss Lipstick FA208 - #203' Rp25.000 di toko Focallure Official Store Tokopedia. Bisa Bebas Ongkir. Iho. <a href="https://t.co/lmC5VbkrG">https://t.co/lmC5VbkrG</a>	positif	@(mention Removal)	Transformation: Remove URL	Regex	Indonesian Stemming	Indonesian Stop word removal
7	#Tokopedia: Ding Dong Deng! 5 hari lagi jangan lupa nonton yaah ENGENE ????	positif	@(mention Removal)	Transformation: Remove URL	Regex	Indonesian Stemming	Indonesian Stop word removal
8	#TokopediaENHYEPEN <a href="https://t.co/FeUGGpsWo">https://t.co/FeUGGpsWo</a>	positif	@(mention Removal)	Transformation: Remove URL	Regex	Indonesian Stemming	Indonesian Stop word removal

**Gambar 4.** Cleansing Data dengan *Gataframework*

Kemudian dilanjutkan tahap pengklasifikasian 1.000 data tweet menggunakan algoritma Adaboost & algoritma *Naïve Bayes*. Maka akan dihasilkan data yang masuk kategori positif atau kategori negatif.



**Gambar 5.** Preprocessing Data Pada Rapid Miner

Dari hasil klasifikasi akan dilanjutkan ke tahap perhitungan akurasi. Tahap akurasi berguna untuk mengetahui tingkat kinerja dari algoritma *Adaboost* dan algoritma *Naïve Bayes*.

### 3.2. Hasil

Hasil dengan algoritma *Naïve Bayes* saja:

**Tabel 2.** Hasil dengan algoritma *Naïve Bayes*

Ket.	accuracy	Precision	Recall	AUC
Persentase	86.80%	23.29%	49.33%	0.511

Hasil dengan algoritma *Naïve Bayes* ditambah fitur SMOTE:

**Tabel 3.** Hasil dengan algoritma *Naïve Bayes* ditambah fitur SMOTE

Ket.	accuracy	Precision	Recall	AUC
Persentase	95.16%	91.29%	100.00%	0.780

Hasil dengan Algoritma *Naïve Bayes & Adaboost* ditambah fitur SMOTE:

**Tabel 4.** Hasil dengan algoritma *Naïve Bayes & Adaboost* ditambah fitur SMOTE

Ket.	accuracy	Precision	Recall	AUC
Persentase	94.95%	90.86%	100.00%	0.950

Dari data hasil uji tersebut data yang menggunakan algoritma *Adaboost & algoritma Naïve Bayes* memiliki kinerja sangat baik dengan akurasi 94.95%, Presisi 90.86%, Recall 100% dan AUC 0,950.

### 4. SIMPULAN

Dari hasil penelitian ini menunjukkan bahwa Algoritma algoritma *Naïve Bayes & Adaboost* jika dioptimasi dengan menggunakan Synthetic Minority Over-sampling Technique (SMOTE) menghasilkan nilai accuracy: 94.95%,



precision: 90.86% , recall: 100.00% dan AUC: 0.950. pada penelitian ini juga menggunakan algoritma *Naive Bayes* dengan optimasi SMOTE menghasilkan accuracy: 95.16%, precision: 91.29%, recall: 100.00% dan AUC: 0.780. sedangkan algoritma *Naive Bayes* tanpa optimasi SMOTE menghasilkan accuracy: 86.80%, precision: 23.29%, recall: 49.33%, dan AUC: 0.511. jadi berdasarkan hasil penelitian ini, dapat disimpulkan bahwa Algoritma *Naive Bayes & Adaboost* yang dioptimasi dengan fitur Synthetic Minority Over-sampling Technique (SMOTE) adalah klasifikasi yang lebih baik digunakan dibandingkan dengan Algoritma *Naive Bayes* dengan optimasi SMOTE maupun Algoritma *Naive Bayes* tanpa optimasi SMOTE.

#### DAFTAR PUSTAKA

- [1] J. Ahmed, "Analisis sentimen dan Klasifikasi Tweet Menggunakan Data Mining," pp. 1471–1474, 2017.
- [2] A. V. Sudiantoro *et al.*, "Analisis Sentimen *Twitter* Menggunakan Text Mining Dengan," vol. 10, no. 2, pp. 398–401, 2018.
- [3] "Text Mining," *Text Mining*. [Online]. Available: [https://id.wikipedia.org/wiki/Penambangan\\_teks](https://id.wikipedia.org/wiki/Penambangan_teks)
- [4] "Sentiment Analysis," *Sentiment Analysis*. [Online]. Available: [https://en.wikipedia.org/wiki/Sentiment\\_analysis](https://en.wikipedia.org/wiki/Sentiment_analysis)
- [5] D. D. Saputra *et al.*, "Optimization Sentiments of Analysis from Tweets in myXLCare using *Naive Bayes* Algorithm and Synthetic Minority over Sampling Technique Method," *J. Phys. Conf. Ser.*, vol. 1471, no. 1, 2020, doi: 10.1088/1742-6596/1471/1/012014.
- [6] M. Shoeb and J. Ahmed, "Sentiment Analysis and Classification of Tweets Using Data Mining," *Int. Res. J. Eng. Technol.*, vol. 4, no. 12, pp. 1471–1474, 2017, [Online]. Available: [www.irjet.net](http://www.irjet.net)
- [7] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di *Twitter* Menggunakan Algoritma Klasifikasi," *J. Teknoinfo*, vol. 14, no. 2, p. 115, 2020, doi: 10.33365/jti.v14i2.679.