

Analisis Sentimen Opini Masyarakat Terhadap Pelayanan BPJS Kesehatan Provinsi Lampung Berbasis Twitter

¹Admi Syarif, ²Arafia Isnayu Akaf, ³Rizky Prabowo, dan ⁴Kurnia Muludi

^{1,2,3,4} Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung
Jalan Soemantri Brojonegoro No.1 Gedung Meneng, Bandar Lampung, Indonesia
e-mail: 1admi.syarif@fmipa.unila.ac.id, 2arafia.isnayu1021@students.unila.ac.id, 3rizky.prabowo@fmipa.unila.ac.id,
4kmuludi@fmipa.unila.ac.id

Abstract — Nowadays, the internet has increased the amount of information stored and accessed through the web at a very fast speed. The internet can be a place to express opinions on health topics, politics, companies, and others. Many social media are used by the people of Lampung in expressing opinions and seeking information. Twitter is one of the communication media that is in great demand by the public. There are various kinds of topics discussed by Twitter users, one of the topics that are currently being discussed is the Lampung Health Social Security Administration Agency (BPJS). Health is also a very important thing and is still a conversation that is often discussed anywhere and anytime. BPJS Health helps the community in overcoming a declining economy, with BPJS Health, the community does not have to pay for medical expenses. Therefore, the service from BPJS Kesehatan Lampung will be carried out by sentiment analysis so that it can be known whether the public opinion about BPJS Kesehatan Lampung is positive or negative. This study uses the Naïve Bayes algorithm. This sentiment uses a dataset from Twitter which uses several keywords regarding BPJS Kesehatan Lampung. Based on the research results, it is known that the Naïve Bayes algorithm has an accuracy value of 89,33%.

Keywords: Naïve Bayes; Sentiment Analysis; Social Security Agency.

1. PENDAHULUAN

Pada era digital ini media sosial digunakan sebagai *platform* untuk berdiskusi, memberikan komentar, berkomunikasi, dan berbagi dalam komunitas yang luas [1]. Ada banyak jenis-jenis media sosial yang digunakan masyarakat untuk berkomunikasi, antara lain Twitter, Facebook, dan Instagram. Media sosial Twitter merupakan salah satu media komunikasi yang banyak diminati oleh masyarakat di seluruh dunia. Twitter adalah *platform microblogging* tempat pengguna membuat *tweet* yang dikirim ke pengikut dan pengguna lain. Pada 2016, Twitter memiliki lebih dari 313 juta pengguna aktif bulanan, termasuk 100 juta pengguna harian. Berdasarkan distribusi secara global, 77% pengguna berasal dari luar Amerika Serikat dan menghasilkan lebih dari 500 juta *tweet* setiap hari [2].

Bidang Kesehatan banyak diperbincangkan di Twitter dan media sosial, contohnya Pelayanan Badan Penyelenggara Jaminan Sosial (BPJS) masyarakat memiliki pemikiran positif atau negatif tentang pelayanan suatu instansi kesehatan. BPJS Kesehatan merupakan substansi yang sah yang mengatur program Jaminan Kesehatan Nasional (JKN) bagi seluruh individu di Indonesia. BPJS Kesehatan memiliki visi dan misi untuk mengakui cakupan medis yang berkualitas dan menawarkan dukungan terbaik kepada masyarakat Indonesia. BPJS Kesehatan dengan Program Jaminan Kesehatan Indonesia Sehat (JKN-KIS) berupaya memperluas cakupan dukungan dengan tujuan agar dapat menjamin setiap individu Indonesia untuk mencapai *Universal Health Coverage* (UHC). Peningkatan jumlah peserta JKN-KIS publik per 30 November 2019 adalah 222.818.475 orang dan tujuannya sebelum akhir tahun 2019 adalah untuk menjangkau mereka semua atau setidaknya 90% dari total penduduk Indonesia. Jumlah anggota pada tahun 2019 meningkat luar biasa dibandingkan dengan informasi dua tahun sebelumnya pada tahun 2017 [3].

Penilaian sejauh mana kualitas pelayanan publik yang diberikan oleh pemerintah yaitu Pelayanan BPJS Kesehatan Provinsi Lampung dilakukan melalui analisis sentimen untuk mengetahui apakah bantuan publik yang diberikan dapat digunakan itu baik atau buruk. Pelayanan BPJS Kesehatan Provinsi Lampung masih

banyak memiliki kekurangan, antara lain belum efektifnya karena kurangnya sosialisasi, sumber daya manusia yang masih kurang, sikap para pelaksana dalam memberikan layanan yang belum terjalin dengan baik yang sangat memungkinkan masyarakat berkomentar mengenai Pelayanan BPJS Kesehatan Provinsi Lampung [4]. Analisis Sentimen juga bisa disebut dengan *opinion mining*. Analisis sentimen adalah teknik untuk mengidentifikasi dan mengenali atau mengkategorikan pendapat pengguna atau untuk layanan seperti lembaga pemerintah, film, dan acara. Sentimen memiliki dua kelas, yaitu positif dan negatif [5]. Metode yang digunakan dalam analisis sentimen adalah *text mining*. *Text mining* adalah analisis teks yang dapat mengekstraksi sumber data dari dokumen untuk menemukan kata yang dapat mewakili isi dokumen dengan dilakukan analisis keterkaitan, keterkaitan, dan kelas antar dokumen. *Text mining* memproses dokumen menjadi klasifikasi teks, ekstraksi informasi, dan ekstraksi kata. Metode ini digunakan untuk mengekstraksi informasi dari sumber data dengan mengidentifikasi dan memeriksa data yang menarik [6].

Pada penelitian ini dilakukan pemodelan topik yang ada pada kelompok sentimen pengguna Twitter terhadap pelayanan dan kebijakan BPJS Kesehatan Lampung menggunakan *Naïve Bayes* agar dapat mengklasifikasi komentar positif atau negatif dari masyarakat umum dan dapat digunakan untuk menilai kualitas layanan sekaligus sebagai bahan pertimbangan kebijakan dari BPJS Kesehatan Provinsi Lampung. Pada penelitian [7] digunakan *Naïve Bayes* untuk klasifikasi sentimen dan menghasilkan 105 data yang memiliki sentimen positif sebanyak 52% dan 95 data ke dalam kategori sentimen negatif sebanyak 48%, menghasilkan nilai akurasi sebesar 86,25%, presisi sebesar 84,92%, *recall* sebesar 87,78%, dan *f1-score* sebesar 86,37% [7].

Algoritme *Naïve Bayes* termasuk bagian *machine learning*. *Machine learning* digunakan pada gagasan menemukan model terbaik untuk data baru dengan menggunakan data yang dikumpulkan sebelumnya [8]. *Naïve Bayes* digunakan untuk mengklasifikasikan teks, kumpulan data pelatihan multidimensi. Beberapa contohnya adalah klasifikasi dokumen, penyaringan *spam*, dan analisis sentimen. Algoritme *Naïve Bayes* dapat membangun model dengan cepat dan memprediksi model dengan cepat dan membutuhkan sedikit pelatihan data. Oleh karena itu, algoritme *Naïve Bayes* disebut “*naïve*” karena kemunculan salah satu fitur dianggap independen dari kemunculan fitur lainnya [9].

Naïve Bayes Classifier memiliki beberapa kelebihan, khususnya cepat dalam komputasi, perhitungan dasar dan presisi tinggi. *Naïve Bayes Classifier* lebih cocok untuk diterapkan pada informasi besar dan dapat menangani data yang tidak lengkap. Meskipun *Naïve Bayes Classifier* memiliki banyak kelebihan, tidak dapat dipungkiri bahwa metode ini juga memiliki kelemahan yaitu probabilitas yang tidak dapat mengukur seberapa besar tingkat keakuratan suatu prediksi. Demikian juga, *Naïve Bayes Classifier* memiliki kekurangan dalam penentuan seleksi atribut sehingga dapat mempengaruhi nilai akurasi [10]. Penelitian ini bertujuan untuk menganalisis komentar masyarakat di media sosial Twitter mengenai Pelayanan Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan Provinsi Lampung guna mengevaluasi kinerja dari Pelayanan BPJS Kesehatan Provinsi Lampung dengan menggunakan algoritme *Naïve Bayes Classifier*.

2. METODOLOGI PENELITIAN

2.1. Pengumpulan Data Teks

Dalam sentimen analisis yang dilakukan memerlukan data yang berasal dari aplikasi *Twitter*. Data yang dibutuhkan merupakan data teks berupa *tweet* yang berkaitan BPJS Kesehatan Lampung. Data teks ini memiliki 2 nilai sentimen, yaitu positif dan negatif. Serta dibuat dengan tabel yang terdiri dari teks dan sentimen. Data yang dikumpulkan sebanyak 650 data teks, setelah itu dilakukan sortir data dan pelabelan manual sehingga data akhir menjadi 300 data teks. Data *tweet* di *Twitter* yang digunakan dalam kurun waktu 3 tahun dari tanggal 30 April 2019 sampai dengan 29 April 2022 dengan menggunakan *keywords* “BPJS Kesehatan Lampung”, “BPJS Lampung”, dan “JKN-KIS Lampung”.

2.2. Pengolahan Data

2.2.1. Pelabelan Manual

Pelabelan data dilakukan secara manual apakah data tersebut memiliki klasifikasi nilai sentimen positif atau negatif. Pelabelan ini dilakukan karena pada data-data yang telah didapatkan belum memiliki nilai sentimen positif atau negatif. Hasil dari pelabelan *manual* yang telah dilakukan yaitu terdapat 267 *tweet* pada kelas positif dan 33 *tweet* pada kelas negatif dan memiliki nilai sebesar 89% pada sentimen positif dan 11% pada sentimen negatif. Gambar 1 menunjukkan sampel pelabelan data yang dilakukan.

	Text	Sentiment
0	sri pedagang tahu yang rasakan manfaat jkn kis...	1
1	ketua komisi informasi lampung alwi agregat da...	1
2	min saya mengajukan pemindahan bpjs dari pbi k...	0
3	bpjs kesehatan bantah peserta jkn kis terlanta...	1
4	halo admin bpjs saya ingin bertanya ibu saya b...	1
...
295	bpjs kesehatan juga jadi syarat bikin sim stnk...	1
296	tabik pun mulai tanggal maret salah satu syara...	1
297	kaa aku lampung jugaa bisa pake bpjs ga yaa kl...	1
298	jangan sakit di lampung ribet banget urusannya...	0
299	bpjs bisa dipake di bidan mandiri yg bekerjasa...	1

300 rows × 2 columns

Gambar 1. Pelabelan manual

2.2.2. Preprocessing Data

Preprocessing data berfungsi untuk membersihkan dan menstandarisasi data *tweet* yang akan diolah dengan menggunakan beberapa metode. Data-data yang terkumpul merupakan data yang masih belum terstruktur dan akan menjadi terstruktur setelah dilakukan proses *preprocessing data*. *Preprocessing data* memiliki beberapa metode yang dapat dilakukan. Beberapa metode tersebut diantaranya *cleaning*, *case folding*, *tokenizing*, dan *stemming*.

2.2.2.1. Cleaning

Cleaning merupakan proses membersihkan informasi teks dari bagian-bagian yang tidak diperlukan atau dianggap sebagai *noise*. Bagian-bagian yang akan diambil antara lain karakter *HTML*, simbol *emoticon*, *hashtag*, *username account*, *retweet*, dan *link URL*. Setelah dilakukan *Cleaning* maka data teks akan bersih dan lebih terstruktur.

2.2.2.2. Case Folding

Case folding merupakan proses standarisasi yang mengubah semua komponen dalam teks menjadi *lowercase*. *Lowercase* yaitu semua huruf pada data teks akan dibuat menjadi huruf kecil semua. Proses ini juga akan menghapus karakter selain huruf yang dianggap tidak berpengaruh terhadap proses pengolahan teks seperti angka dan tanda baca.

2.2.2.3. Tokenizing

Tokenizing merupakan proses yang pemisahan kalimat dalam data teks menjadi sebuah pecahan kata tunggal atau token yang terpisah dan terstruktur.

2.2.2.4. Stopword Removal

Stopword removal merupakan sebuah proses dalam pemilihan kata dengan menghilangkan kata yang tidak berguna dalam data yang didapatkan sehingga dapat meningkatkan hasil akurasi yang lebih optimal dalam klasifikasi tersebut.

2.2.2.5. Stemming

Stemming merupakan cara paling umum untuk mengubah kata-kata yang baru-baru ini ditambahkan ke dalam struktur kata dasar mereka sesuai dengan strukturnya. Maka dapat diasumsikan bahwa kata yang diubah memiliki arti dan makna yang sama dengan kata dasarnya.

2.2.3. Pembagian Data

Pembagian data dilakukan dengan cara membagi data menjadi dua yaitu meliputi data latih dan data uji. Data latih dan uji berfungsi agar data yang diolah dapat dilakukan secara optimal. Proses ini menggunakan *4-Fold Cross Validation* yang dapat dilihat pada Tabel 1.

Tabel 1. Pembagian data latih dan data uji

<i>Fold</i>	Data Latih	Data Uji
1	75-299	0-74
2	0-74, 150-299	75-149
3	0-149, 225-299	150-224
4	0-224	225-299

Pembagian data dilakukan menggunakan *Library sklearn.model_selection*. Pembagian data menggunakan 75% data latih dan 25% data uji dari 300 data teks yang berasal dari Twitter. Data uji yang digunakan 75 data teks pada tiap *fold*. Kode Program untuk pembagian data adalah dapat dilihat pada kode program berikut.

```
#Data Fold ke - 1
X_train_1 = tweets.loc[75:299, "Text"]
Y_train_1 = tweets.loc[75:299, "Sentiment"]
X_test_1 = tweets.loc[0:74, "Text"]
Y_test_1 = tweets.loc[0:74, "Sentiment"]

#Data Fold ke - 2
X_train_2 = pd.concat([tweets.loc[0:74, "Text"], tweets.loc[150:299, "Text"]],
ignore_index=True)
Y_train_2 = pd.concat([tweets.loc[0:74, "Sentiment"], tweets.loc[150:299,
"Sentiment"]], ignore_index=True)
X_test_2 = tweets.loc[75:149, "Text"]
Y_test_2 = tweets.loc[75:149, "Sentiment"]
```

```
#Data Fold ke - 3
X_train_3 = pd.concat([tweets.loc[0:149, "Text"], tweets.loc[225:299, "Text"]],
ignore_index=True)
Y_train_3 = pd.concat([tweets.loc[0:149, "Sentiment"], tweets.loc[225:299,
"Sentiment"]], ignore_index=True)
X_test_3 = tweets.loc[150:224, "Text"]
Y_test_3 = tweets.loc[150:224, "Sentiment"]

#Data Fold ke - 4
X_train_4 = tweets.loc[0:224, "Text"]
Y_train_4 = tweets.loc[0:224, "Sentiment"]
X_test_4 = tweets.loc[225:299, "Text"]
Y_test_4 = tweets.loc[225:299, "Sentiment"]
```

2.3. Klasifikasi Sentimen

Klasifikasi sentimen dilakukan dengan menggunakan algoritme *Naive Bayes*. Selanjutnya akan didapatkan hasil klasifikasi berupa *confusion matrix* yang dapat digunakan untuk mendapatkan nilai akurasi, presisi, *recall*, dan *f1-score*.

2.3.1. *Naive Bayes*

Naive Bayes Classifier adalah pengklasifikasi statistik yang dapat digunakan untuk memprediksi probabilitas menjadi anggota suatu kelas. *Naive Bayes Classifier* didasarkan pada teorema *Bayes*, yang memiliki kemampuan klasifikasi yang sama dengan pohon keputusan. *Naive Bayes Classifier* telah terbukti memberikan akurasi dan kecepatan tinggi saat diterapkan pada *database* yang berisi data dalam jumlah besar. *Naive Bayes Classifier* adalah algoritma yang populer dan bagus untuk data berdimensi tinggi seperti teks [11].

Perhitungan probabilitas dalam *Naive Bayes* menggunakan pendekatan algoritme *Bayes* menggunakan persamaan (1) [10].

$$P(H | X) = \frac{P(X | H) P(H)}{P(X)} \quad (1)$$

Keterangan:

- X = Data dengan kelas yang belum diketahui;
- H = Hipotesis data X merupakan suatu kelas spesifik;
- $P(H|X)$ = Probabilitas hipotesis H berdasarkan kondisi X ;
- $P(X|H)$ = Probabilitas hipotesis X berdasarkan kondisi H ;
- $P(H)$ = Probabilitas hipotesis H ;
- $P(X)$ = Probabilitas hipotesis X

Algoritme *Naive Bayes* dapat dilakukan dengan beberapa tahap. Tahap yang dilakukan antara lain *crawling data* dan *preprocessing data*. *Crawling data* dilakukan dengan cara memberikan *keywords* atau kata kunci, serta dilakukan proses *labeling data* untuk melabelkan sebuah data teks dengan beberapa sentimen. Tahap selanjutnya yang dilakukan yaitu *preprocessing data* untuk menyeleksi data lalu mengubahnya menjadi data yang terstruktur. Pada tahap ini hal-hal yang dilakukan yaitu melakukan proses *cleaning* untuk mengurangi *noise* dan *stopword removal* untuk menghapus kata-kata yang tidak memiliki makna. Proses *tokenization*

digunakan dalam mengidentifikasi kata-kata dan memisahkan beberapa kalimat menjadi sebuah istilah berdasarkan spasi dan tanda baca. Tahap terakhir yang dilakukan dalam *preprocessing* yaitu *stemming*, dengan mengubah kata imbuhan menjadi kata dasar [12].

2.4. Evaluasi *Confusion Matrix*

Confusion matrix adalah tabel yang digunakan untuk menggambarkan kinerja model klasifikasi pada set data pengujian yang nilainya telah diketahui. *Confusion matrix* sendiri relatif sederhana dan mudah untuk mengerti. Cara kerja *confusion matrix* cukup sederhana, yaitu dengan membagi hasil klasifikasi dari algoritme *machine learning*, Tabel *confusion matrix* dapat dilihat pada Tabel 2.

Tabel 2. *Confusion matrix*

Fakta	Prediksi	
	Negatif	Positif
Negatif	TN (<i>True Negative</i>)	FP (<i>False Positive</i>)
Positif	FN (<i>False Negative</i>)	TP (<i>True Positive</i>)

Keterangan:

- *True Positive* (TP) adalah data positif yang terdeteksi positif;
- *False Negative* (FN) adalah data positif yang terdeteksi negatif;
- *True Negative* (TN) adalah jumlah data negatif yang terdeteksi negatif;
- *False Positif* (FP) adalah data negatif namun terdeteksi positif.

Tabel *confusion matrix* digunakan untuk mengukur kinerja suatu metode klasifikasi dengan menghitung nilai akurasi (*accuracy*), presisi (*precision*), *recall*, dan *f1-score* [13].

a. Akurasi

Akurasi (*accuracy*) merupakan rasio prediksi benar (positif dan negatif) dengan keseluruhan data.

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2)$$

b. Presisi

Presisi (*precision*) merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif.

$$\text{Presisi} = \frac{TP}{TP+FP} \times 100\% \quad (3)$$

c. *Recall*

Recall merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif.

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (4)$$

d. *F1-score*

F1-Score adalah perbandingan rata-rata presisi dan *recall* yang dibobotkan.

$$F1 - score = 2 \times \frac{Recall \times Precision}{Recall+Precision} \quad (5)$$

3. HASIL DAN PEMBAHASAN

3.1. Pengujian berdasarkan Kombinasi Pembagian Data Latih dan Data Uji

Tahap pengujian yang pertama yaitu melakukan pembagian data dengan *4-fold* dengan menggunakan algoritme *Naïve Bayes*. Algoritme yang dilakukan yaitu dengan menggunakan pemodelan *Multinomial Naïve Bayes* dapat dilihat pada Tabel 3.

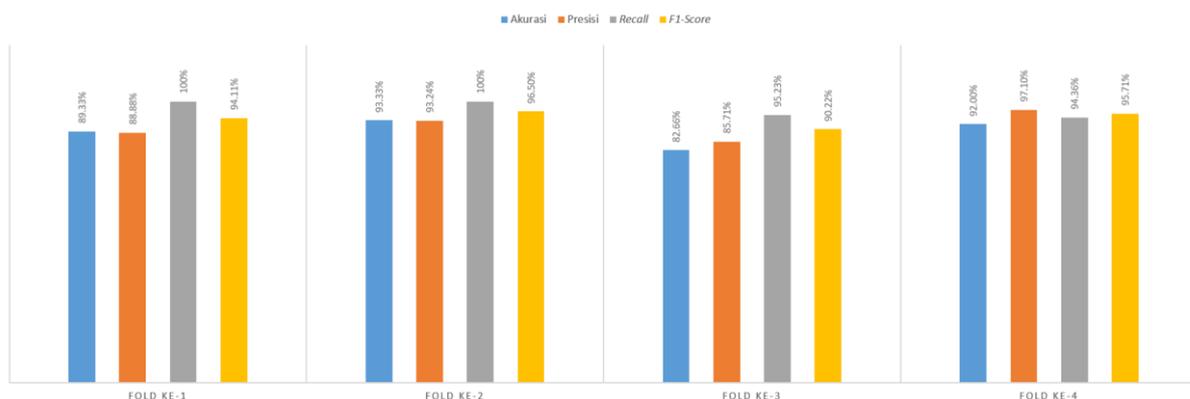
Tabel 3. Hasil pengujian *naïve bayes*

<i>Fold</i>	TN	FN	FP	TP	Akurasi (%)	Presisi(%)	<i>Recall</i> (%)	<i>F1-score</i> (%)
1	3	0	8	64	89,33	88,88	100	94,11
2	1	0	5	69	93,33	93,24	100	96,50
3	2	3	10	60	82,66	85,71	95,23	90,22
4	2	4	2	67	92,00	97,10	94,36	95,71
Rata-rata					89,33	91,23	97,39	94,13

Pada Tabel 3 maka dapat disimpulkan bahwa hasil dari perhitungan klasifikasi dengan menggunakan algoritme *Naïve Bayes* menggunakan pembagian 75% data latih dan 25% data uji menghasilkan rata-rata akurasi sebesar 89,33%, presisi sebesar 91,23%, *recall* sebesar 97,39%, dan *f1-score* sebesar 94,13%.

3.2. Hasil Diagram Akurasi, Presisi, *Recall*, dan *F1-Score*

Pada pengujian ini didapatkan nilai akurasi, presisi, *recall*, dan *f1-score* yang didapatkan dari pengujian menggunakan *4-fold cross validation* dengan 75% data latih dan 25% data uji. Terdapat juga grafik perbandingan masing-masing *fold* dari akurasi, presisi, *recall*, dan *f1-score*. Diagram dari akurasi dapat dilihat pada Gambar 2.



Gambar 2. Akurasi, presisi, recall, dan f1-score

Pada Gambar 2 dapat dilihat bahwa *fold* ke-1 memiliki nilai akurasi sebesar 89,33%, presisi sebesar 88,88%, *recall* sebesar 100%, dan *f1-score* sebesar 94,11%. *Fold* ke-2 memiliki nilai akurasi sebesar 93,33%, presisi sebesar 93,24%, *recall* sebesar 100%, dan *f1-score* sebesar 96,50%. *Fold* ke-3 memiliki nilai akurasi sebesar 82,66%, presisi sebesar 85,71%, *recall* sebesar 95,23%, dan *f1-score* sebesar 90,22%. *Fold* ke-4 memiliki nilai akurasi sebesar 92,00%, presisi sebesar 97,10%, *recall* sebesar 94,36%, dan *f1-score* sebesar 95,71%.

4. KESIMPULAN

Penelitian ini telah berhasil mengaplikasikan algoritme *Naïve Bayes* dalam melakukan sentimen analisis opini masyarakat pengguna Twitter terhadap pelayanan Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan Provinsi Lampung. Penelitian ini menggunakan data dari Twitter dengan *keywords* BPJS Kesehatan Lampung, BPJS Lampung, dan JKN-KIS Lampung dalam kurun waktu 3 tahun, dari tanggal 30 April 2019 sampai dengan 29 April 2022. Hasil dari eksperimen cenderung bernilai positif. Hasil *tweet* dari pelabelan manual memiliki 267 *tweet* positif dan 33 *tweet* negatif dengan persentase *tweet* positif sebanyak 89% dan persentase *tweet* negatif sebanyak 11%. Berdasarkan hasil pengujian menggunakan *4-fold cross validation* dengan pengujian data latih 75% dan data uji 25% dari algoritme *Naïve Bayes*, dapat diketahui bahwa *fold* ke-1 memiliki nilai akurasi sebesar 89,33%, presisi sebesar 88,88%, *recall* sebesar 100%, dan *f1-score* sebesar 94,11%. *Fold* ke-2 memiliki nilai akurasi sebesar 93,33%, presisi sebesar 93,24%, *recall* sebesar 100%, dan *f1-score* sebesar 96,50%. *Fold* ke-3 memiliki nilai akurasi sebesar 82,66%, presisi sebesar 85,71%, *recall* sebesar 95,23%, dan *f1-score* sebesar 90,22%. *Fold* ke-4 memiliki nilai akurasi sebesar 92,00%, presisi sebesar 97,10%, *recall* sebesar 94,36%, dan *f1-score* sebesar 95,71%.

DAFTAR PUSTAKA

- [1] S. A. Putri, P. D. Kusuma, & C. Setianingsih, "Clustering Topik Pada Data Sentimen Bpjs Kesehatan Menggunakan Metode Latent Dirichlet Allocation (lda)," *eProceedings of engineering*, vol. 8, no. 5, pp. 6097–6105, 2021.
- [2] D. Zimbra, A. Abbasi, D. Zeng, & H. Chen, "The state-of-the-art in twitter sentiment analysis: A review and benchmark evaluation," *ACM Trans. Manag. Inf. Syst.*, vol. 9, no. 2, 2018, doi: 10.1145/3185045.
- [3] M. Nurmalasari, N. Anggita Temesvari, S. N. Maula, "Analisis Sentimen terhadap Opini Masyarakat dalam Penggunaan Mobile-JKN untuk Pelayanan BPJS Kesehatan Tahun 2019," *Heal. Inf. Manag. J. ISSN*, vol. 8, no. 1, pp. 2655–9129, 2020.
- [4] A. Gusnita, "Analisis Implementasi Jaminan Sosial Nasional Pada Badan Penyelenggara Jaminan Sosial Kesehatan (Bpjs) Kota Metro," *Pros. SNaPP2017 Sos. Ekon. dan Hum.*, pp. 692–698, 2017.
- [5] P. Mehta & S. Pandya, "A review on sentiment analysis methodologies, practices and applications," *Int. J. Sci. Technol. Res.*, vol. 9, no. 2, pp. 601–609, 2020.
- [6] R. Novendri, A. S. Callista, D. N. Pratama, & C. E. Puspita, "Sentiment Analysis of YouTube Movie Trailer Comments Using Naïve Bayes," *Bull. Comput. Sci. Electr. Eng.*, vol. 1, no. 1, pp. 26–32, 2020, doi: 10.25008/bcsee.v1i1.5.
- [7] M. Yunus, M. Husni, & M. M. Mufadhhal, "Klasifikasi Sentimen Terhadap Badan Penyelenggara Jaminan Sosial (BPJS) Pada Media Sosial Twitter Menggunakan Naive Bayes," *Smatika J.*, vol. 11, no. 02, pp. 81–91, 2021, doi: 10.32664/smatika.v11i02.577.
- [8] U. Ö. Osmanoglu, O. N. Atak, K. Çağlar, H. Kayhan, & T. Can, "Sentiment Analysis for Distance Education Course Materials: A Machine Learning Approach," *J. Educ. Technol. Online Learn.*, vol. 3, no. 1, pp. 31–48, 2020, doi: 10.31681/jetol.663733.
- [9] M. Abbas, K. Ali, A. Jamali, K. A. Memon, & A. A. Jamali, "Multinomial Naive Bayes Classification Model for Sentiment Analysis Classification for Sentiment Analysis View project

Antenna Designing View project Multinomial Naive Bayes Classification Model for Sentiment Analysis,” *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, vol. 19, no. 3, p. 62, 2019, doi: 10.13140/RG.2.2.30021.40169.

- [10] H. Muhamad, C. A. Prasajo, N. A. Sugianto, L. Surtiningsih, & I. Cholissodin, “Optimasi Naïve Bayes Classifier Dengan Menggunakan Particle Swarm Optimization Pada Data Iris,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 3, p. 180, 2017, doi: 10.25126/jtiik.201743251.
- [11] I. E. Tiffani, “Optimization of Naïve Bayes Classifier By Implemented Unigram, Bigram, Trigram for Sentiment Analysis of Hotel Review,” *J. Soft Comput. Explor.*, vol. 1, no. 1, pp. 1–7, 2020, doi: 10.52465/josce.v1i1.4.
- [12] W. A. Prabowo & C. Wiguna, “Sistem Informasi UMKM Bengkel Berbasis Web Menggunakan Metode SCRUM,” *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 149, 2021, doi: 10.30865/mib.v5i1.2604.
- [13] K. Muludi, M. Surya Akbar, D. A. Shofiana & A. Syarif, “Sentiment Analysis Of Energy Independence Tweets Using Simple Recurrent Neural Network,” *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 15, no. 4, pp. 339–348, 2021, doi: 10.22146/ijccs.xxxx.