

Analisis Sentimen Opini Publik Mengenai Wajib BPJS Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes

Muhammad Rakha^{1*}

¹Fakultas Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Jl. Raya Puspiptek No. 46, Kel. Buaran, Kec. Serpong, Kota Tangerang Selatan. Banten 15310, Indonesia

Email: 1*muhammadrakha004@gmail.com

(* : coressponding author)

Abstrak– Pada tanggal 6 Januari 2022, Presiden Joko Widodo telah menandatangani Instruksi Presiden (Inpres) kepada publik. Salah satu poin Inpres Nomor 1 Tahun 2022 tentang BPJS sebagai kewajiban pelayanan publik, menjadi perbincangan publik, mengenai pemohon SIM, STNK, dan Surat Keterangan Catatan Kepolisian (SKCK) yang harus menyertai keanggotaan BPJS Kesehatan. Dikarenakan banyaknya keluhan masyarakat mengenai pedoman pendidikan kepresidenan, maka penelitian ini dilakukan untuk menilai persepsi masyarakat dengan menggunakan algoritma Naive Bayes dengan menggunakan seleksi fitur TF-IDF. Penelitian ini menggunakan 362 data yang dibagi menjadi dua kelas yaitu positif dan negatif. Proses klasifikasi memberikan hasil akurasi terbaik dengan persentase 88,89% menggunakan pemilihan fungsi TF-IDF dengan data latih 90% dan data uji 10%.

Kata Kunci: Teks Mining, Sentimen Analisis, Naïve Bayes, BPJS Kesehatan

Abstract– On January 6, 2022, President Joko Widodo signed the Presidential Instruction (Inpres) to the public. One of the points of Presidential Instruction Number 1 of 2022 concerning BPJS as a public service obligation, is a public discussion, concerning applicants for SIMs, STNK marks, and Police Certificates (SKCK) that must accompany BPJS Health membership. Due to the many public complaints about presidential education guidelines, this study was conducted to assess public perceptions using the Naive Bayes algorithm using TF-IDF feature selection. This study uses 362 data which is divided into two classes, positive and negative. The classification process gives the best accuracy results with a percentage of 88.89% using the selection of the TF-IDF function with 90% training data and 10% test data.

Keywords: Text Mining, Sentiment Analysis, Naïve Bayes, BPJS Kesehatan

1. PENDAHULUAN

Pada tanggal 6 Januari 2022, Presiden Joko Widodo menandatangani Intruksi Presiden (Inpres) nomor 1 tahun 2022 tentang Optimalisasi Pelaksanaan Program Jaminan Kesehatan Nasional yang mana berisi intruksi kepada Kementrian dan Pemerintah Daerah untuk menunjukkan kepesertaan BPJS Kesehatan sebagai syarat untuk pelayanan kepada masyarakat.

Beberapa poin dari Inpres no 1 tahun 2022 tentang BPJS menjadi syarat layanan publik yang menjadi perbincangan di kalangan masyarakat, di antaranya pemohon SIM, STNK, dan Surat Keterangan Catatan Kepolisian (SKCK) harus melampirkan kepesertaan BPJS Kesehatan, intruksi juga ditunjukkan untuk calon Jemaah Umrah dan Jemaah Haji khusus untuk aktif menjadi peserta BPJS Kesehatan, penerima Kredit Usaha Rakyat (KUR) harus menjadi peserta BPJS Kesehatan, dan terakhir pemohon pendaftaran peralihan hak tanah karena jual beli merupakan peserta aktif BPJS Kesehatan.

Kebijakan BPJS jadi Syarat pelayanan publik tentunya menimbulkan dampak positif dan negatif. Pemberitaan ini menjadi perdebatan pro dan kontra dikalangan masyarakat terkait Kebijakan Wajib BPJS, hal ini dikarenakan pengesahan Kebijakan Wajib BPJS dinilai terlalu cepat dan masih ada beberapa pasal yang menurut masyarakat kurang tepat sasaran dan tidak sesuai dengan kejadian di lapangan.

Maka dari itu perlu diketahui opini dari masyarakat Indonesia mengenai Kebijakana Wajib BPJS. Biasanya masyarakat mengemukakan pendapatnya melalui sosial media. Salah satu aplikasi yang sering digunakan masyarakat Indonesia adalah Twitter. Twitter banyak digunakan orang untuk menyampaikan keluh kesahnya mulai dari keluh kesah mengenai kehidupan sehari-hari ataupun keluh kesah terhadap layanan yang diberikan baik dari pemerintah atau bidang lainnya. Oleh karena itu sangat efisien jika menggunakan Twitter sebagai media untuk mengambil data mengenai keluhan masyarakat Indonesia terkait kebijakan Wajib BPJS (Fariz, Suhardono, & Verdiana, 2021)..

Kebijakan BPJS jadi Syarat pelayanan publik tentunya menimbulkan dampak positif dan negatif. Pemberitaan ini menjadi perdebatan pro dan kontra dikalangan masyarakat terkait Kebijakan Wajib BPJS, hal ini dikarenakan pengesahan Kebijakan Wajib BPJS dinilai terlalu cepat dan masih ada beberapa pasal yang menurut masyarakat kurang tepat sasaran dan tidak sesuai dengan kejadian di lapangan.

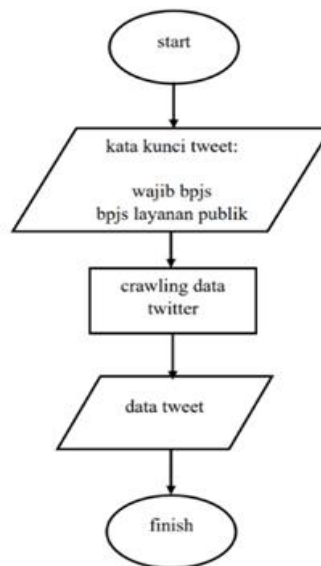
Maka dari itu perlu diketahui opini dari masyarakat Indonesia mengenai Kebijakan Wajib BPJS. Biasanya masyarakat mengemukakan pendapatnya melalui sosial media. Salah satu aplikasi yang sering digunakan masyarakat Indonesia adalah *Twitter*. *Twitter* banyak digunakan orang untuk menyampaikan keluh kesahnya mulai dari keluh kesah mengenai kehidupan sehari-hari ataupun keluh kesah terhadap layanan yang diberikan baik dari pemerintah atau bidang lainnya. Oleh karena itu sangat efisien jika menggunakan *Twitter* sebagai media untuk mengambil data mengenai keluhan masyarakat Indonesia terkait kebijakan Wajib BPJS (Fariz, Suhardono, & Verdiana, 2021).

Berdasarkan uraian di atas, maka akan dilakukan penelitian dengan judul “ANALISIS SENTIMEN OPINI PUBLIK MENGENAI WAJIB BPJS PADA MEDIA SOSIAL *TWITTER* MENGGUNAKAN METODE *NAÏVE BAYES*”.

2. METODOLOGI PENELITIAN

2.1 Teknik Pengumpulan data

Pengumpulan data pada penelitian ini dilakukan dengan cara crawling. Proses crawling dilakukan dengan menggunakan Rapid Miner. Crawling tweet dilakukan secara otomatis. Kata yang dicari adalah “wajib bpjs” dan “bpjs layanan publik”. Diagram alur proses crawling yaitu sebagai berikut:

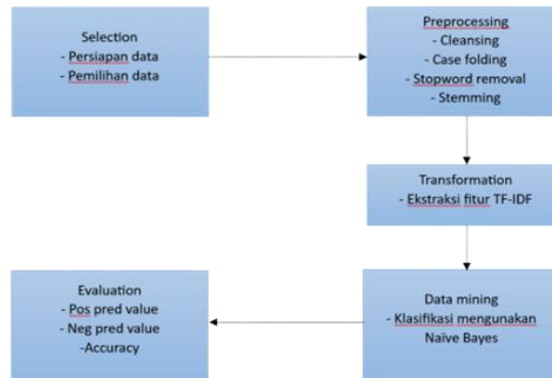


Gambar 1. Digram Alur Proses *Crawling Dataset*

sPada gambar 1 dapat dijelaskan bahwa untuk melakukan pengambilan data menggunakan crawling pada RapidMiner, kata kunci tweet yang akan digunakan harus dimasukkan terlebih dahulu kedalam tools dengan rentang waktu yang dibutuhkan sehingga menghasilkan data tweet yang siap digunakan (Fikri, Sabrila, & Azhar, 2020).

2.2 Metode Penelitian

Metode penelitian yang akan digunakan dalam penelitian ini adalah *Knowledge Discovery in Database (KDD)*. *Knowledge Discovery in Database (KDD)* merupakan sebuah proses menemukan informasi yang berguna dalam database. Proses KDD adalah proses yang mengidentifikasi *pattern* dalam data yang benar, unik, berguna, dan dimengerti (Juliansa, 2019). Adapun tahapan dalam *Knowledge Discovery in Database (KDD)* sebagai berikut:



Gambar 2. Alur Penelitian

Pada gambar 2 dapat dijelaskan tahapan dalam pengolahan data yang akan digunakan dalam penelitian ini yaitu:

- a. Seleksi Data (*Data Selection*).
- b. Pemilihan Data (*Preprocessing*).
- c. Transformasi Data (*Data Transformation*).
- d. *Data Mining*.
- e. Evaluasi/Interpretasi.

3. ANALISA DAN PEMBAHASAN

Hasil penelitian merupakan tahapan dari penelitian yang telah dilakukan berupa hasil klasifikasi sentimen positif dan sentimen negatif terhadap Wajib BPJS pada *twitter* menggunakan *Naive Bayes*.

3.1 Selection

Data sentimen terhadap Wajib BPJS pada *twitter* memasuki tahap *selection*. Tahap seleksi dilakukan oleh peneliti. Dari 601 data *tweet* didapatkan 362 *tweets*. Berikut ini merupakan data sentimen yang telah di seleksi oleh peneliti:

tweet	kelas
1 @BatuKepla @gandi_sinta @alisyarief Ahh sudah lah bong.....	Negatif
2 Mau jual tanah, wajib BPJS. Mau dapat BLT, wajib vaksin wal...	Negatif
3 @txtdrpemerintah skrg. bbm naik, minyak gada, apa2 wajib ...	Negatif
4 RT @FaridYiyit: Siap2 rumah dan barang pribadi lainnya di b...	Negatif
5 @HariKusdiyanto @TeddGus @jokowi @YaqutCQoumas @...	Negatif
6 @kesehatan_bersama BPJS memberikan akses kepada peser...	Positif
7 @_AnakKolong @PBIDI Banyak dokter benci pada pemerin...	Negatif
8 RT @satia_nagara: @papa_loren Bp cape juga harta+asset k...	Negatif
9 RT @ImamSantoso01: @Mdy_Asmara1701 Susah jadi rakyat...	Negatif
10 @papa_loren Bp cape juga harta+asset kekayaan semakin b...	Negatif
11 @Mdy_Asmara1701 Susah jadi rakyat, jangan dipikir gampa...	Negatif
12 ya sama aja kayak lo gak minta dilahirin di indo tapi otomati...	Negatif
13 @_Sridiana_3va Kalau RT kek gitu, mending di tawur saja b...	Negatif
14 Jd gini alurnya: Askes swasta, dokter ngeresepin obat paten ...	Negatif
15 @Beritasatu Indonesia lebih baik, khususnya yg berada di il...	Negatif
16 RT @detikfinance: Kepesertaan BPJS Kesehatan menjadi sala...	Negatif
17 @lolokKAN BPJS membantu masyarakat untuk lebih fokus p...	Positif
18 @tempodotco Bagusan demo tu pemerintah,gak becus me...	Negatif
19 @susipudjiastuti Wow.. Padahal nilai investasinya gede bgt. ...	Negatif
20 @Chanaicoklat @arifnovianto_id @sicepat_ekspres Setauku ...	Negatif
21 @democrazymedia Baik bagi opung, krna dpt banyak, bany...	Negatif
22 RT @caritautdotcom: Kritik Kebijakan Wajib Punya BPJS Kese...	Negatif
23 Kritik Kebijakan Wajib Punya BPJS Kesehatan, Ombudsman ...	Negatif

Gambar 3. Data Hasil Seleksi

Pada gambar 2 dapat dijelaskan bahwa data yang berhasil diseleksi oleh peneliti dibagi kedalam dua kelas yaitu kelas sentimen positif dan kelas sentimen negatif dengan jumlah data sebanyak 362 data *tweet* sentimen yang terdiri dari 27 *tweet* dengan sentimen positif dan 335 data *tweet* dengan sentimen negatif.

3.2 Preprocessing

Data sentimen yang telah melalui tahap *selection*, kemudian dilakukan tahap *preprocessing* untuk menghilangkan kata yang kurang berpengaruh dalam proses pengklasifikasian. Berikut merupakan hasil dari *preprocessing* data yang akan dilakukan.

a. Cleansing

```
<<SimpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 5

[1] kritik kebijakan wajib Punya BPJS Kesehatan Ombudsman RI Jangan Terburuburu
[2] Ombudsman Republik Indonesia meminta Pemerintah tidak terburuburu dalam memberlakukan kepesertaan BPJS Kesehatan sebagai syarat mengakses pelayanan publik
[3] adanya syarat wajib BPJS utk pengurusan dokumen banyak keluhan di persulitnya mengaktifkan BPJS orang mau patuh kan baik mohon petugas BPJS utk keluar dari meja kerja datang antrian layani kasih solusi lengkap biar nggak bolak x
[4] mohon di bantu wong cilik masa simstnk wajib ikut BPJS soalnya cari buat makan saja bisa untung kalau di kk ada daftar BPJS harus semua te rus bayarnya pakai daun dan kalau buat berobat kayak anak tiri di rumah sakit
[5] sudah swab mbak tapi ga usahlah gapapa sekarang menang musimnya Musin segala sesuatu yang ga jelas ya dok batuk kenaikan BBM wajib BPJS sa npe urusan sambo
```

Gambar 4. Hasil Proses *Cleansing*

Pada gambar 4 dapat dijelaskan bahwa data yang sudah dilakukan proses *cleansing* berhasil menghilangkan *mention*, *hashtag*, dan tanda baca (*punctuation*) yang tidak berpengaruh.

b. Case Folding

```
<<SimpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 5

[1] kritik kebijakan wajib punya bpjs kesehatan ombudsman ri jangan terburuburu
[2] ombudsman republik indonesia meminta pemerintah tidak terburuburu dalam memberlakukan kepesertaan bpjs kesehatan sebagai syarat mengakses pelayanan publik
[3] adanya syarat wajib bpjs utk pengurusan dokumen banyak keluhan di persulitnya mengaktifkan bpjs orang mau patuh kan baik mohon petugas bpjs utk keluar dari meja kerja datang antrian layani kasih solusi lengkap biar nggak bolak x
[4] mohon di bantu wong cilik masa simstnk wajib ikut bpjs soalnya cari buat makan saja bisa untung kalau di kk ada daftar bpjs harus semua te rus bayarnya pakai daun dan kalau buat berobat kayak anak tiri di rumah sakit
[5] sudah swab mbak tapi ga usahlah gapapa sekarang menang musimnya musin segala sesuatu yang ga jelas ya dok batuk kenaikan bbm wajib bpjs sa npe urusan sambo
```

Gambar 5. Hasil Proses *Case Folding*

Pada gambar 5 dapat dijelaskan bahwa data yang sudah dilakukan proses *case folding* berhasil mengubah huruf kapital menjadi huruf kecil

c. Stopword Removal

```
<<SimpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 5

[1] kritik kebijakan wajib bpjs kesehatan ombudsman ri terburuburu
[2] ombudsman republik indonesia pemerintah terburuburu memberlakukan kepesertaan bpjs kesehatan syarat mengakses pelayanan publik
[3] syarat wajib bpjs utk pengurusan dokumen keluhan persulitnya mengaktifkan bpjs orang patuh mohon petugas bpjs utk meja kerja data ngot antrian layani kasih solusi lengkap biar nggak bolak x
[4] mohon bantu cilik simstnk wajib bpjs cari makan untung kk daftar bpjs bayarnya pakai daun berobat kayak anak tiri rumah sakit
[5] swab mbak ga usahlah gapapa musimnya musin ga ya dok batuk kenaikan bbm wajib bpjs sampe urusan sambo
```

Gambar 6. Hasil Proses *Stopword Removal*

Dari gambar 6 dapat dijelaskan bahwa data yang telah melalui tahap *stopword removal* telah dibandingkan dengan kosa kata yang terdapat dalam kamus *stopword*. *Stopword* yang digunakan dalam setiap studi kasus dapat disesuaikan sesuai kebutuhan data penelitian.

d. Stemming

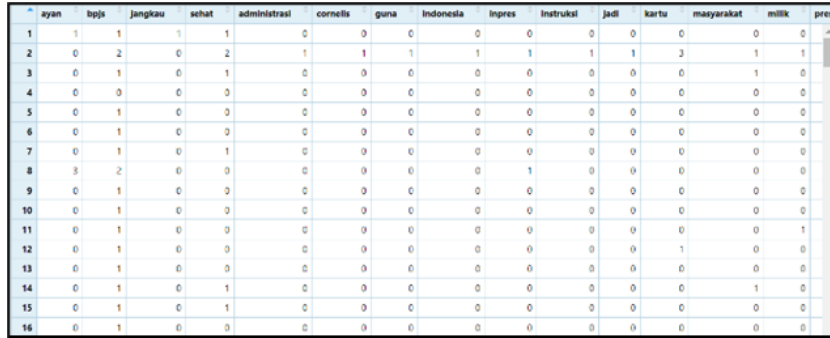
```
<<SimpleCorpus>>
Metadata: corpus specific: 1, document level (indexed): 0
Content: documents: 5

[1] kritik bijak wajib bpjs sehat ombudsman r terburuburu
[2] ombudsman republik indonesia pemerintah terburuburu memberlakukan kepesertaan bpjs sehat syarat akses ayan publik
[3] syarat wajib bpjs utk urus dokumen keluh sulit aktif bpjs orang patuh mohon tugas bpjs utk meja kerja datang antri layan kasih solusi Ten gkap biar nggak bolak x
[4] mohon bantu cilik simstnk wajib bpjs cari makan untung kk daftar bpjs bayar pakai daun obat kayak anak tiri rumah sakit
[5] swab mbak ga usah gapapa musim musim ga ya dok batuk naik bbm wajib bpjs sampe urusan sambo
```

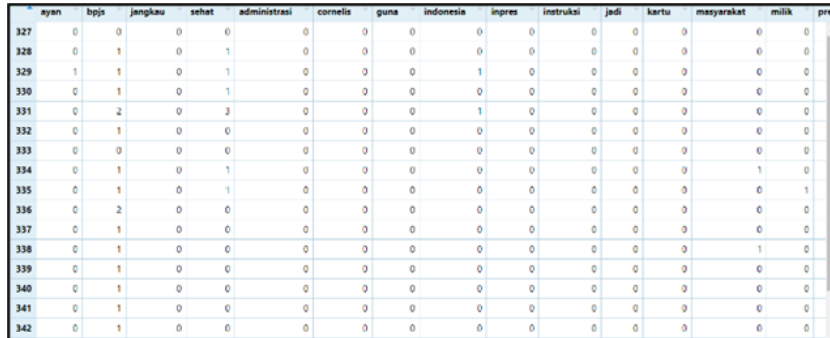
Gambar 7. Hasil Proses *Stemming*

Dari gambar 7 dapat dijelaskan bahwa data *tweet* yang telah melalui tahap *stemming* merupakan data bersih yang sudah dihilangkan imbuhanannya dan menjadi kata dasar sehingga dapat dilanjutkan ke tahap selanjutnya.

3.3 Transformation



Gambar 8. Bobot Pada Data Latih



Gambar 9. Bobot Pada Data Uji

Proses pembobotan dilakukan dengan cara membandingkan bobot setiap kata pada data uji dengan kata pada data latih. Jika bobot probabilitas negatif lebih tinggi, maka hasil tweet tersebut adalah kategori negatif. Namun, jika bobot probabilitas positifnya lebih tinggi, maka hasil tweet tersebut berkategori positif.

3.4 Data Mining

Pada tahap *data mining* adalah tahap dilakukannya klasifikasi data teks menggunakan *Naïve Bayes*. Dilakukan pembagian data menjadi data latih dan data uji dengan menggunakan persentase tertentu sebagai rasio pembagiannya.

Dalam penelitian ini, dilakukan lima kali pengujian dengan persentase pembagian data latih dan data uji yang berbeda-beda. Pengujian pertama dilakukan dengan menggunakan persentase 50% data latih dan 50% data uji, pengujian kedua menggunakan persentase 60% data latih dan 40% data uji, pengujian ketiga menggunakan persentase 70% data latih dan 30% data uji, pengujian keempat menggunakan persentase 80% data latih dan 20% data uji, dan pengujian kelima menggunakan persentase 90% data latih dan 10% data uji (Prana, Indriati, & Adikara, 2019).

Tabel 1. Pembagian Data Latih dan Data Uji

Persentase (%)	Data Latih	Data Uji
50% dan 50%	181	181
60% dan 40%	217	145
70% dan 30%	253	109
80% dan 20%	289	73
90% dan 10%	326	36

Berikut adalah langkah-langkah dalam proses pelatihan untuk mengklasifikasikan data latih secara manual yang telah melalui proses *preprocessing*, yaitu:

Tabel 2. Data *Training*

<i>Tweet</i>	Kalimat	Label
<i>Tweet 1</i>	proses jual beli tanah lampir bpjs sehat status aktif	negatif
<i>Tweet 2</i>	bodoh coba panjang stnk wajib pake bpjs sehat cook aneh	negatif
<i>Tweet 3</i>	urus simstnk wajib pakai bpjs sehat gimana	negatif
<i>Tweet 4</i>	wajib bpjs bantu kurang beban biaya sehat masyarakat	positif
<i>Tweet 5</i>	bpjs mudah kontrol biaya sehat atur uang	positif
<i>Tweet 6</i>	bpjs bantu masyarakat waspada sehat keluarga	positif

Berdasarkan tabel 2, maka dapat dibuat sebuah model probabilitas sebagai berikut:

$$P(V_j) = \frac{docs_j}{training} = \frac{3}{6} = 0,5$$

$$P(a_{proses}|V_{negatif}) = \frac{1 + 1}{21 + 32} = 0,037$$

$$P(a_{proses}|V_{positif}) = \frac{0 + 1}{21 + 32} = 0,021$$

$$P(a_{jual}|V_{negatif}) = \frac{1 + 1}{21 + 32} = 0,037$$

$$P(a_{jual}|V_{positif}) = \frac{0 + 1}{21 + 32} = 0,021$$

$$P(a_{beli}|V_{negatif}) = \frac{1 + 1}{21 + 32} = 0,037$$

$$P(a_{beli}|V_{positif}) = \frac{0 + 1}{21 + 32} = 0,021$$

Probabilitas setiap kata pada data latih dapat dilihat pada tabel berikut ini:

Tabel 3. Perhitungan Probabilitas Data *Training*

<i>Class</i>	$P(V_j)$	$P(a_i V_j)$								
		proses	jual	beli	tanah	lampir	bpjs	sehat	status	aktif
negatif	0,5	0,037	0,037	0,037	0,037	0,037	0,075	0,075	0,037	0,037
positif	0,5	0,021	0,021	0,021	0,021	0,021	0,086	0,086	0,021	0,021

Perhitungan probabilitas pada table 3, digunakan sebagai model yang akan menjadi acuan untuk menentukan data uji. Proses pengujian berikut akan menggunakan perhitungan probabilitas dan bertujuan untuk menemukan probabilitas tertinggi (Wulandari, Enri, & Primajaya, 2020).

Tabel 4. Data *Testing*

<i>Tweet</i>	Kalimat	Kelas
<i>Tweet 7</i>	jual beli tanah wajib lampir bpjs amat negara zalim rakyat	?
<i>Tweet 8</i>	bpjs program pemerintah bantu tingkat sejahtera masyarakat	?
<i>Tweet 9</i>	bpjs bantu masyarakat fokus kembang karier khawatir biaya sehat	?

Tweet 7:

$$= P(a_{jual}/V_{negatif}) \times P(a_{beli}/V_{negatif}) \times P(a_{tanah}/V_{negatif}) \times P(a_{wajib}/V_{negatif}) \times P(a_{lampir}/V_{negatif}) \times P(a_{bpjs}/V_{negatif}) \times P(a_{amat}/V_{negatif}) \times P(a_{negara}/V_{negatif}) \times P(a_{zalim}/V_{negatif}) \times P(a_{rakyat}/V_{negatif}) \times P(V_j)$$

$$= 0,037 \times 0,037 \times 0,037 \times 0,056 \times 0,037 \times 0,075 \times 0,5 = 0,0000000039357381$$

Tweet 7:

$$= P(a_{jual}/V_{positif}) \times P(a_{beli}/V_{positif}) \times P(a_{tanah}/V_{positif}) \times P(a_{wajib}/V_{positif}) \times P(a_{lampir}/V_{positif}) \times P(a_{bpjs}/V_{positif}) \times P(a_{amat}/V_{positif}) \times P(a_{negara}/V_{positif}) \times P(a_{zalim}/V_{positif}) \times P(a_{rakyat}/V_{positif}) \times P(V_j)$$

$$= 0,021 \times 0,021 \times 0,021 \times 0,043 \times 0,021 \times 0,086 \times 0,5 = 0,0000000003595953$$

Berdasarkan perhitungan yang telah dilakukan, didapatkan hasil sebagai berikut:

Tabel 5. Hasil Perhitungan Probabilitas Data *Testing*

<i>Tweet</i>	Nilai Probabilitas	
	Negatif	Positif
<i>Tweet 7</i>	0,0000000039357381	0,0000000003595953
<i>Tweet 8</i>	0,0000121500000000	0,0001816750000000
<i>Tweet 9</i>	0,0000000164025000	0,0000010155632500

3.5 Evaluation

Berikut merupakan hasil akurasi terhadap klasifikasi sentimen Wajib BPJS, yaitu sebagai berikut:

Tabel 6. Tabel Akurasi Model Klasifikasi

Rasio <i>Split</i>	Accuracy (%)
50:50	79,56%
60:40	77,24%
70:30	82,57%
80:20	86,30%
90:10	88,89%

Berdasarkan hasil seleksi yang dilakukan oleh peneliti, data sentimen terhadap Wajib BPJS berjumlah 362 *tweet* yang terdiri dari 335 *tweet* negatif dan 27 *tweet* positif. Klasifikasi data sentimen terhadap Wajib BPJS pada *twitter* menggunakan *Naïve Bayes* dilakukan dengan lima pengujian dan mendapatkan akurasi tertinggi pada rasio *split* 90:10 sebesar 88,89% dengan pembagian data yakni 36 data uji.

Secara keseluruhan dapat disimpulkan bahwa algoritma *Naïve Bayes* pada penelitian ini menghasilkan akurasi yang tinggi. Hal ini dibuktikan juga oleh (Utami & Erfina, 2022) dalam jurnalnya yang melakukan klasifikasi dengan menggunakan *Naïve Bayes* mendapatkan hasil akurasi 90,71% dari 75 ulasan wisatawan yang memberikan ulasan positif 69 dan 9 wisatawan memberikan ulasan negatif.

4. KESIMPULAN

Berdasarkan hasil penelitian mengenai Analisis Sentimen ini, maka penulis dapat menyimpulkan sebagai berikut:

- Algoritma *Naive Bayes* dapat dimanfaatkan untuk analisis sentimen mengenai Wajib BPJS pada *social media twitter* dilakukan dengan membagi data yang telah melewati tahap *preprocessing* menjadi data latih dan data uji. Pengujian dilakukan sebanyak 5 kali dengan presentase 50% data latih dan 50% data uji, 60% data latih dan 40% data uji, 70% data latih dan 30% data uji, 80% data latih dan 20% data uji, 90% data latih

dan 10% data uji. Pada presentase 90% data latih dan 10% data uji hasil klasifikasi 326 data latih dan 36 data uji menggunakan *Naïve Bayes* dengan seleksi fitur TF-IDF pada RStudio didapatkan 28 data kelas negatif diprediksi benar sebagai kelas negatif, 0 data kelas positif diprediksi sebagai kelas negatif, 4 data kelas negatif diprediksi sebagai kelas positif, dan 4 data kelas positif diprediksi benar sebagai kelas positif.

- b. Penerapan algoritma *Naïve Bayes* untuk klasifikasi data sentimen terhadap Wajib BPJS yang dimana data bersumber dari sosial media *twitter* dengan jumlah data sebanyak 362 *tweet*. Hasil akurasi terbaik merupakan klasifikasi menggunakan *Naïve Bayes* dengan seleksi fitur TF-IDF yaitu sebesar 88,89% dengan presentase data sebanyak 90% data latih dan 10% data uji.

REFERENCES

- Fariz, T. R., Suhardono, S., & Verdiana, S. (2021). Pemanfaatan Data Twitter Dalam Penanggulangan. *Cogito Smart Journal*, 135-147.
- Fikri, M. I., Sabrila, T. S., & Azhar, Y. (2020). Perbandingan Metode Naive Bayes dan Support Vector Machine pada Analisis Sentimen Twitter. *SMATIKA Jurnal*, Volume 10 Nomor 02 71-76.
- Juliansa, H. (2019). DATA MINING ROUGH SET DALAM MENGANALISA KINERJA DOSEN STMIK BINA NUSANTARA JAYA LUBUKLINGGAU. *JUSIM (Jurnal Sistem Informasi Musirawas)*, 11-17.
- Prana, P. A., Indriati, & Adikara, P. P. (2019). Klasifikasi Komentar Body Shaming Beauty Vlogger Pada Youtube Menggunakan Metode BM 25 dan K-Nearest Neighbour. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 7328-7334.
- Utami, D. S., & Erfina, A. (2022). Analisis Sentimen Objek Wisata Bali Di Google Maps Menggunakan Algoritma Naive Bayes. *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 418-427.
- Wulandari, P. D., Enri, U., & Primajaya, A. (2020). Analisis Sentimen Terhadap Kenaikan Iuran. *JURNAL ILMU KOMPUTER DAN TEKNOLOGI*, 18-22.