

Analisis Sentimen Pada Tweets Divisi Humas Polri Dengan Metode *Naïve Bayes Classifier*

Caldiyastovan Mohi¹, Haditsah Annur², Roys Pakaya³

Ilmu Komputer ,Teknik Informatika, Universitas Ichsan Gorontalo, Gorontalo, Indonesia
Email: ¹Caldiyastovan Mohi@gmail.com

Abstrak - Berita merupakan cerita atau keterangan tentang suatu kejadian atau peristiwa yang hangat dibicarakan. Hal ini dapat dimanfaatkan media online untuk memberikan berita terupdate kepada masyarakat. Penggunaan media sosial dikalangan masyarakat dalam menyebarkan informasi dengan cepat sehingga jarak dan waktu tidak menjadi halangan bagi pengguna media sosial untuk dapat menerima dan mengakses perkembangan informasi secara luas tanpa terkendala. Twitter Merupakan salah satu platform media sosial yang banyak digunakan oleh individu, pemerintah dan organisasi, termasuk kepolisian, untuk mempublikasikan berita tentang instansi terkait. Dengan menerapkan salah satu fungsi dari *Text mining* yaitu *Naïve Bayes Classifier* untuk menganalisis sentimen masyarakat terhadap *tweet* dari akun *twitter* Divisi Humas Polri. menganalisis sentimen pada *tweet* secara manual tidak lagi efektif, maka diperlukan metode *Naïve Bayes Classifier* yang dapat secara otomatis menganalisis sentimen pada *tweet* menjadi opini positif dan negatif. Makadengan menggunakan algoritma ini penulis mendapatkan nilai akurasi yang cukup tinggi. Pada penelitian ini penulis mendapatkan nilai akurasi yang tinggi sebesar 76% pada saat melakukan *splitting* sebesar 10% data *testing* dan 90% data *training*. Tetapi pada saat melakukan *preprocessing* dan implemetasi data baru pada *framework streamlit* membutuhkan waktu hingga 1 menit untuk dapat memproses data.

Kata Kunci: *Naïve Bayes*, Humas Polri, twitter, analisis sentimen

Abstract *News is a story or information about an event that is hotly discussed. It can be utilized by online media to provide updated news to the public. The use of social media among the public in disseminating information is fast so that distance and time are not an obstacle for social media users to be able to receive and access information developments widely without constraints. Twitter is a social media platform that is widely used by individuals, governments, and organizations, including the police, to publish news about related agencies. By applying one of the functions of text mining, namely the Naïve Bayes Classifier to analyze public sentiment towards tweets from the Indonesian Police Public Relations Division's Twitter account, manually analyzing sentiment on tweets is no longer effective, so a Naïve Bayes Classifier method is needed to automatically analyze sentiment on tweets into positive and negative opinions. So, by using this algorithm, this study gets a fairly high accuracy value. In this study, a high accuracy value of 76% when splitting 10% of testing data and 90% of training data. But when preprocessing and implementing new data in the streamlit framework, it takes up to 1 minute to process the data.*

Keywords: *Naïve Bayes*, Police Public Relations, Twitter, sentiment analysis

1. PENDAHULUAN

Berita merupakan cerita atau keterangan tentang suatu kejadian atau peristiwa yang hangat dibicarakan. Berdasarkan hasil riset yang dilakukan oleh lembaga global GFK dan Indonesia Digital Association (IDA) dilaksanakan di lima kota besar di Indonesia sepanjang tahun 2015, presentase konsumsi berita melalui televisi sebesar 91 persen, surat kabar 31 persen dan radio sebesar 15 persen dari 1.521 responden, Angka terendah dibandingkan dengan konsumsi berita melalui online mencapai 96 persen. Hal ini dapat dimanfaatkan media online untuk memberikan berita terupdate kepada masyarakat. [1]

Divisi Humas Polri merupakan unsur pengawas dan pembantu pimpinan dibidang humas di tingkat mabas Polri. Divisi humas polri memiliki akun media sosial salah satunya di *twitter*. Pengguna *twitter* dapat mengetahui informasi terkini terkait kegiatan polri melalui akun Divisi Humas Polri. Divisi Humas Polri memiliki akun *twitter* yang dimana setiap orang yang memiliki keprihatinan atau saran dapat mempostingnya.

Twitter Merupakan salah satu platform media sosial yang banyak digunakan oleh individu, pemerintah, dan organisasi, termasuk kepolisian, untuk mempublikasikan berita tentang instansi terkait. Dengan cara ini seseorang dapat mengikuti perkembangan terkini dan menyuarakan pendapat mereka. Sehingga sentimen publik terhadap suatu topik dapat dipelajari dan menjadi sumber informasi dari media sosial.[3]

Analisis sentimen merupakan analisa yang dilakukan pada suatu peristiwa dari pendapat yang berdasarkan pada sikap seseorang tentang suatu objek. Analisis sentimen biasanya melakukan pengumpulan opini untuk mengetahui opini masyarakat dalam postingan di sosial media salah satunya adalah *twitter*. [4] Mengetahui opini publik terhadap suatu objek untuk keperluan tugas akhir maka membutuhkan Analisis sentimen. Ada dua kategori opini publik yaitu negatif dan positif tergantung pada pandangan publik tentang suatu objek tersebut. Sehingga penulis membutuhkan suatu analisis terhadap opini-opini publik dengan tujuan penelitian agar bisa dijadikan tolok ukur baik atau tidaknya persepsi masyarakat tentang polri.

2. TINJAUAN PUSTAKA

2.1 Twitter

Berisi *Twitter* adalah platform media sosial dengan layanan microblogging yang memungkinkan penggunaannya untuk menulis sesuatu, yang dikenal sebagai *tweet*. Ini banyak digunakan oleh masyarakat umum. [7] Penyebaran berita resmi dari berbagai akun kantor berita maupun opini publik melalui *twitter*. Jack dorsey membuat *twitter* pada tahun 2006 dengan alamat <http://www.twitter.com>. Untuk menggunakan *twitter*, anda harus membuat akun terlebih dahulu, dan anda bisa mendaftar di alamat tersebut. *Twitter* memfasilitasi penyebaran informasi yang lebih cepat, yang kemudian akan menjadi bahan diskusi di antara pengguna *twitter*. Keunggulan *twitter* antara lain memiliki jumlah pengikut yang tidak terbatas dan tidak ada timbal balik; misalnya, pengguna A dapat mengikuti akun *twitter* pengguna B meskipun pengguna B tidak mengikuti akun pengguna A. *Twitter* juga banyak digunakan sebagai sumber data untuk penambangan teks. Ini karena pencarian *twitter* juga bisa berdasarkan pesan status tertentu selain nama akun. Data Mining. Berikut ini adalah istilah pada *twitter* :

- a. *Tweet* adalah status yang ditulis di kotak yang sudah disediakan *twitter* dan mencakup informasi, gambar, trek, opini, dan ulasan. *Tweet* juga disebut pesan. Ada batasan karakter 280 untuk pesan atau informasi yang ingin ditulis.
- b. *Mention* Digunakan untuk memanggil pengguna *twitter* lain pada sebuah *tweet*. *Mention* dituliskan dengan '@' diikuti dengan nama pengguna lain.
- c. Like merupakan salah satu icon yang digunakan untuk menandai *tweet* yang di sukai
- d. *Direct message* Mengirimkan pesan secara langsung ke orang lain tanpa ada yang mengetahuinya.
- e. *Replay* atau balasan bisa digunakan untuk membalas *tweet* pengguna lain.
- f. *Retweet* adalah penanda apabila kita setuju dengan isi dari *tweet* orang lain kemudian dibagikan ke pengguna *tweet* lain.
- g. *Hashtag* dapat digunakan untuk mengidentifikasi *tweet* tertentu di *twitter*. Hastag "#" digunakan bersamaan dengan nama topik. Hastag juga meningkatkan visibilitas setiap *tweet* individu.
- h. *Trending topic* Pembicara yang sedang trending karena sedang ramai dibicarakan oleh pengguna *twitter*.

2.2. Twitter API

Dalam *Twitter* memperkenalkan API yang tangguh, atau antarmuka pemrograman aplikasi, yang memungkinkan banyak pengguna yang mengakses data *twitter*. Data *twitter* dapat digunakan untuk berbagai keperluan, termasuk mencari informasi populer, mendapatkan mention, dan mendapatkan Retweet dari akun di *twitter* untuk mendongkrak pengembangan aplikasi. [4] API (Application Programming Interface) adalah sekumpulan perintah, fungsi dan protokol yang dapat digunakan untuk membangun perangkat lunak pada sistem operasi tertentu. Ini juga berfungsi sebagai dokumentasi pembuatan perangkat lunak dengan antarmuka, fungsi, kelas dan struktur. Pemrograman dapat menghubungkan satu perangkat lunak ke perangkat lunak lainnya menggunakan API.

2.3. Crawling Data

Dalam proses pengambilan data atau *Crawling* data dari *twitter* dimana data tersebut untuk keperluan tugas akhir peneliti. *Crawling* data ini bisa dilakukan dengan cara membuat program dengan

memasukkan kata kunci yang kita inginkan untuk mencari *tweet*. Misalnya, “@divHumas_Polri” pada saat memasukan kata kunci berupa *mention* user name pada program yang kemudian akan di proses untuk mengambil *tweet* yang *mention* ke akun Divisi Humas Polri. Berikut adalah code program python untuk proses *crawling*.

```
for i,tweet in enumerate(sntwitter.TwitterSearchScrapper('@DivHumas_Polri since:2022-01-01
until:2022-10-28').get_items()):
    if i > 1000:
        break
    tweets_list.append([tweet.date, tweet.id, tweet.content, tweet.user.username,
        analize_sentiment(tweet.content)])
```

Gambar 1. Code Program Crawling

2.3. Preprocessing

Preprocessing dilakukan setelah mempersiapkan dataset. Pada tahapan *preprocessing* terdapat beberapa proses yang akan dilakukan sehingga data menjadi terstruktur. Berikut ini tahap dalam proses *preprocessing*.

1. *Case folding* merupakan pengubahan huruf besar menjadi huruf kecil pada semua ulasan yang ada pada dokumen data *training* dan data *testing*. Hanya huruf “a” sampai dengan huruf “z” yang diterima dan karakter selain huruf akan dihilangkan.[8]

Tabel 1. Case Folding

Sebelum <i>case folding</i>	Hasil <i>case folding</i>
PAK PRESIDEN @jokowi TOLONG USAT TUNTAS KASUS SUPIR DI PAPUA YANG DIBUNUH GEROMBOLAN ORANG BERSIFAT SETAN!	pak presiden jokowi tolong usat tuntas kasus supir di papua yang dibunuh gerombolan orang bersifat setan

2. *Tokenization* adalah pemisahan kalimat (*string*) yang setiap kalimatnya disusun menjadi potongan tunggal.[8]

Tabel 2. Tokenization

Sebelum <i>Tokenization</i>	Hasil <i>Tokenization</i>
polisi harus jujur dan santun karna kebanyakan boong dan arogan	Polisi Harus Jujur dan santun karna kebanyakan boong dan arogan

3. *Stopword* adalah proses penghilang kata yang tidak berkaitan dengan topik dokumen, karena kata tersebut tidak akan mempengaruhi keakuratan dalam klasifikasi sentimen.[8]

Tabel 3. Stopword

Sebelum Stopword	Hasil Stopword
polisi harus jujur dan santun karena kebanyakan bohong dan arogan	polisi harus jujur dan santun kebanyakan bohong dan arogan

4. *Stemming* adalah proses untuk mengubah kata yang terdapat dalam suatu dokumen ke kata akarnya (*root word*) yang terdapat dalam sistem IR.[8]

Tabel 2. Steaming

Sebelum Stemming					Hasil Stemming				
polisi	harus	jujur	dan	Santun	polisi	harus	jujur	dan	Santun
karena	kebanyakan	bohong	dan	arogan	karena	banyak	bohong	dan	arogan

2.3. Analisis Sentimen

Studi komputasi tentang penilaian, pendapat, dan perasaan orang tentang suatu entitas atau peristiwa dikenal sebagai analisis sentiment. Analisis sentimen adalah metode untuk memproses data teks secara otomatis untuk mengekstrak data dan memahami opini

2.4. Naïve Bayes Classifier

Naïve Bayes Classifier adalah teknik klasifikasi statistik yang memperhitungkan probabilitas kemungkinan kelas, termasuk probabilitas signifikansi kelas.

Yang perlu dilakukan adalah menghitung nilai probabilitas masing-masing dengan persamaan (1).

$$P(W_{ij} | V_{ij}) = \frac{n_{ij}+1}{n+[kosakata]} \dots\dots\dots(1)$$

Keterangan :

$P(W_{ij} | V_{ij})$ = Peluang Kategori j Ketikas Munculnya Kata i

$n_{ij} + 1$ = Frekuensi Munculnya Kata Pada Sebuah Kategori

n = Jumlah Semua Kata Pada Suatu Kategori Dalam Dokumen

$[kosakata]$ = Jumlah Seluruh Kata (Distinc) Pada Seluruh Data Latih

Setelah itu, menghitung probabilitas kategori dengan menggunakan persamaan (2).

$$P(V_{ij}) = \frac{[docsj]}{[contoh]} \dots\dots\dots(2)$$

Keterangan :

$P(V_{ij})$ = Probabilitas Dokumen Kategori

$[docsj]$ = Jumlah Semua Dokumen Pada Sebuah Kategori

$[contoh]$ = Jumlah Total Data Yang Dilatih

Untuk proses pengklasifikasian, data yang dimasukan belum diketahui kategorinya. Peroses ini dilakukan pencarian kata pada data yang dimasukan sesuai dengan data latih. Untuk klasifikasi data yang dimasukan menggunakan persamaan (3).

$$VMAP = \operatorname{argmax}_{v_{ij} \in v} P(V_{ij}) \prod_i P(a_i | V_{ij}) \dots\dots\dots(3)$$

Fase pengujian dan pelatihan implementasi *Naïve Bayes* terdiri dari dua proses, yaitu proses *testing* dan *training*. Proses latihan digunakan untuk membangun model analisis yang dapat digunakan untuk menganalisis dalam jumlah besar sebagai pengklasifikasian data mentah yang baru.

2.5. Pembobotan Kata

Pembobotan kata adalah suatu fitur yang dibutuhkan untuk transformasi data sehingga teks digital bisa diubah menjadi model yang efisien agar proses analisis siap untuk dijalankan. Bobot *term* dapat dianalogikan sebagai data biner (true and false), frekuensi, dan frekuensi invers dokumennya (TF-IDF).[6].Metode TF-IDF digunakan untuk mengetahui seberapa sering kata yang muncul didalam dokumen.[12] TF-IDF dapat dilakukan dengan beberapa cara sebagaai berikut.

$$IDF = \log \frac{D}{DF} \dots\dots\dots(4)$$

$$TF - IDF = tf * idf \dots\dots\dots(5)$$

Keterangan :

- D : Jumlah Dokumen Dalam Dataset
- DF : Jumlah Dokumen Yang Mengandung Kata Tersebut
- Tf : *Term Frequency*
- Idf : *Inversi Document Frequency*

TF-IDF adalah metode yang sering digunakan karena mudah untuk dipahami, efisien dan hasil yang akurat.

2.6. Python

Python merupakan bahasa pemrograman yang banyak digunakan perusahaan besar maupun developer untuk mengembangkan berbagai macam aplikasi berbasis desktop, web, dan mobile. peneliti menggunakan *Python* yang mempunyai sintaksis kode yang sangat jelas dan juga dilengkapi fungsional yang tinggi sehingga mempermudah peneliti dalam menyelesaikan tugas akhir yang terkait dengan pemrograman.[13]

2.7. Confusion Matrix

Confusion Matrix digunakan untuk mengevaluasi sistem klasifikasi yang berdasarkan kemampuan parameter prediktif suatu model. parameter prediktif adalah parameter yang mengukur kualitas set uji statistik yang diberikan berdasarkan atribut yang ada dalam sampel. *Confusion Matrix* dapat dilihat pada tabel berikut.

Tabel 3. confusion matrix

Actual Class	Predicted Class	
	Kelas=Positif	Kelas=Negatif
Kelas = Positif	TP(True Positive)	FN(False Negatif)
Kelas = Negatif	FP(False Positive)	TN(True Negatif)

$$akurasi = \frac{(TP+TN)}{TP+TN+FP+FN} \times 100 \dots\dots\dots (6)$$

$$precision = \frac{TP}{TP+FP} \times 100 \dots\dots\dots (7)$$

$$recall = \frac{TP}{TP+FN} \times 100 \dots\dots\dots (8)$$

$$f - measure = \frac{(2 \times recall \times presisi)}{(recall + presisi)} \dots\dots\dots (9)$$

Keterangan :

TP(*True Positive*) : jumlah data yang memiliki kategori positif diklasifikasikan positif.

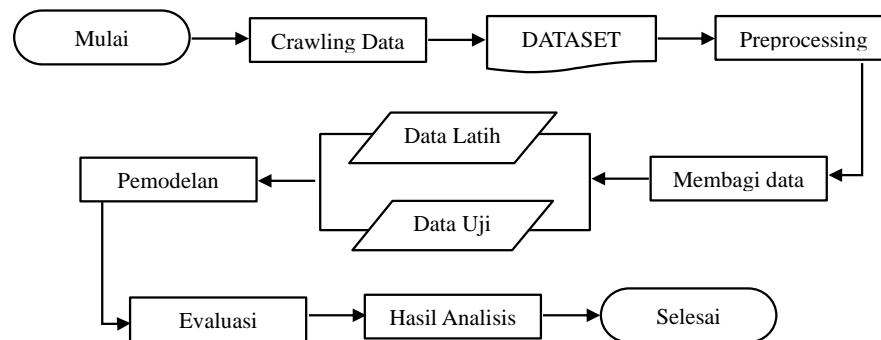
TN(*True Negatif*) : jumlah data yang memiliki kategori negatif diklasifikasikan negatif.

FN(*False Negatif*) : jumlah data yang memiliki kategori positif diklasifikasikan negatif.

TN(*True Negatif*) : jumlah data yang memiliki kategori negatif diklasifikasikan positif.

3. METODE PENELITIAN

Model yang diusulkan dapat digambarkan menggunakan flowchart berikut ini :



Gambar 2. Pemodelan Yang diusulkan

3.1 Preprocessing

Tahapan setelah dataset yaitu *Preprocessing* data. *Preprocessing* data adalah tahap pengolahan data mentah menjadi data yang akan diproses pada tahapan selanjutnya. Ada beberapa tahapan *Preprocessing* yang akan dilakukan yaitu merubah huruf besar menjadi huruf kecil dan menghilangkan seluruh tanda baca (*case folding*), memisahkan sekumpulan karakter dalam teks menjadi suatu kata (*tokenization*), menghilangkan karakter-karakter atau kata yang tidak digunakan (*filtering stopword*), dan merubah kata berimbuhan menjadi kata dasar (*stemming*).

3.2 Naive Bayes Classifier

Pada tahap ini dilakukan klasifikasi sentimen dengan menggunakan metode *Naive Bayes Classifier*. Menyiapkan data yang telah diberi label sebelumnya untuk digunakan dalam proses klasifikasi (*data training*), merupakan langkah awal dalam proses metode klasifikasi *Naive Bayes*. Penggunaan library *python* akan memproses penetapan sentimen ke jumlah data yang disiapkan sebelumnya. Peneliti juga akan secara manual memberi label data dengan membagi tanggapan menjadi kategori positif dan negatif.

3.3 Evaluasi

Evaluasi dilakukan untuk memverifikasi hasil dari klasifikasi dengan melakukan pengukuran keakuratan pada metode klasifikasi. Tahap evaluasi ini menggunakan metode *confusion matrix* untuk pembandingan hasil dari klasifikasi yang dilakukan pada sistem sesuai dengan hasil yang sebenarnya.

4. HASIL DAN PEMBAHASAN

Pada penelitian ini menggunakan dataset *tweet* berbahasa indonesia yang diambil dari akun *twitter* Divisi Humas Polri. Kemudian peneliti melakukan pemrosesan dan pengklasifikasian pada dataset berdasarkan pemodelan yang di usulkan. Jika terjadi outlier atau *tweet* yang tidak berhubungan dengan topik yang akan diteliti dan tidak memenuhi kriteria untuk dijadikan dataset pada proses pengumpulan dataset, maka peneliti akan menyeleksi *tweet* tersebut dari dataset agar tidak mengganggu pada saat proses klasifikasi.

Tabel 6. Labeling Dataset

No	text	label
1	Lestarkan Semangat Persatuan & Kesatuan	Positif
2	Terima kasih atas responnya, Ndan poldajateng Kami tunggu perkembangan kasus tersebut! 🙏	Positif
3	Dibikin ga masuk akal biar susah dan milih buat nembak aja, mayan dapet cuan tuhh 😊	Negatif
4	Mantap Pak ListyoSigitP	Positif
5	Tilang saja sesuai aturan ...!! jgn buat2 aturan dewe☐	Positif
6	Delik pidana pencemaran nama baik RI1 sudah terbukti nie,. harusnya tetap bisa proses hukum makhluk satu itu.	Positif
7	Tidak laku klo gini	Negatif
8	dan jajaran kepolisian lainnya bisa segera ditangkap pelakunya, sebelum kepercayaan masyarakat hilang, sehingga akan dihakimi massal.	Negatif
9	Saya juga termasuk salah satu yang lolos, kemudian saya pernah ditanya apa keluhannya saat melakukan tes lapangan tsb, dan kemudian saya menyatakan hal sama seperti ini kepada polisi tsb, namun jawabn yg diberikn sngt tdk memuaskan.	Negatif
10	Setuju. Ayo donk pak tunjukkan kalau Polri tidak berpihak, gerakan para buzzer ini masif dan organized lho. Tindak mereka pak.	Positif
.....
122	Ini anak sudah kurang ajar Marwah Presiden jokowi seenaknya dia samakan dengan Tukan Obat pinggir jalan... Tindak... 🤔	Positif

4.1 Klasifikasi Naïve Bayes

Splitting data adalah tahapan yang dilakukan untuk membagi dataset menjadi data *testing* dan data *training*, proses *splitting* dilakukan sebelum proses klasifikasi. Berikut adalah source code dari proses *splitting* data.

```

162
163     t_size = st.radio(
164         "Pilih pembagian splitting data",
165         (0.1,0.2,0.3))
166     st.write("size splitting :", t_size)
167     if st.button('Mulai klasifikasi'):
168         st.success("3. naïve Bayes")
169         st.write("Kernel GaussianNB")
170         #vectorizer = TfidfVectorizer(max_features=2500)
171         from sklearn.naive_bayes import GaussianNB
172         model_g = GaussianNB()
173         from sklearn.model_selection import train_test_split
174         v_data = vectorizer.fit_transform(dataset.values()).toarray()
175         X_train, X_test, y_train, y_test = train_test_split(v_data, dnb['label'], test_size=t_size, random_state=0)

```

Selanjutnya proses klasifikasi, peneliti menggunakan *library Scikit-Learn* pada streamlit untuk proses klasifikasi *Naïve Bayes*.

```

167     if st.button('Mulai klasifikasi'):
168         st.success("3. naïve Bayes")
169         st.write("Kernel GaussianNB")
170         #vectorizer = TfidfVectorizer(max_features=2500)
171         from sklearn.naive_bayes import GaussianNB
172         model_g = GaussianNB()
173         from sklearn.model_selection import train_test_split
174         v_data = vectorizer.fit_transform(dataset.values()).toarray()
175         X_train, X_test, y_train, y_test = train_test_split(v_data, dnb['label'], test_size=t_size, random_state=0)
176         nbmodel = model_g.fit(X_train, y_train)
177         nbmodel.class_count_
178         from sklearn.metrics import accuracy_score
179         y_preds = model_g.predict(X_test)
180         predict = pd.array(y_preds)
181         true_label = pd.array(y_test)
182         st.write("Data Testing Sebenarnya")
183         st.write(true_label)
184         st.write("Data Testing Vg Diprediksi")
185         st.write(predict)
186         proba = nbmodel.predict_proba(X_test)
187         st.write("Nilai Probabilitas")
188         st.write(proba)
189         st.write('nilai akurasi adalah ',
190         accuracy_score(y_test, y_preds)*100)
191

```

Gambar 4.2 : Code Klasifikasi

kemudian peneliti melakukan perhitungan untuk mendapatkan hasil klasifikasi berdasarkan model yang di bangun. Pada perhitungan ini peneliti menggunakan *microsoft excel*. Untuk total data *training* menggunakan 10 dokumen dari hasil perhitungan TF-IDF dengan jumlah 6 dokumen sentimen positif dan 4 dokumen sentimen negatif. Berikut ini adalah tabel hasil akhir dari perhitungan klasifikasi.

PROBABILITAS NAÏVE BAYES	
PROBABILITAS	Sentimen
0,5923958044	POSITIF
0,4000000000	NEGATIF
HASIL AKHIR KLASIFIKASI	
POSITIF	0,59

4.2 Evaluasi

Tahapan terakhir evaluasi model yaitu melakukan perhitungan manual *confusion matrix* untuk mengetahui nilai akurasi, *presicion*, *recall* dan *f1-score* dari klasifikasi tersebut. pada tabel confusion matrix diberi keterangan FP(False Positif), FN(False Negatif), TP(True Positif) dan TN(True Negatif). Berikut adalah tabel perhitungan manual *confusion matrix*.

Selanjutnya melakukan proses perhitungan manual dengan menggunakan rumus berikut ini.

1. Akurasi

Untuk menghitung akurasi ini dilakukan dengan menggunakan persamaan (10) berikut

$$\begin{aligned} \text{akurasi} &= \frac{(TP+TN)}{TP+TN+FP+FN} \times 100 & (10) \\ &= \frac{(8+2)}{8+2+3+0} \\ &= \frac{10}{13} = 0,76 \times 100 = 76\% \end{aligned}$$

2. Precision

Untuk menghitung Precision ini dilakukan dengan menggunakan persamaan berikut.(11)

$$\begin{aligned} \text{Presisi} &= \frac{(TP)}{FP+TP} & (11) \\ &= \frac{8}{3+8} = \frac{8}{11} = 0,73 \end{aligned}$$

3. Recall

Untuk menghitung *Recall* ini dilakukan dengan menggunakan persamaan (12) berikut.

$$\begin{aligned} \text{Recall} &= \frac{TP}{FN+TP} & (12) \\ &= \frac{8}{0+8} = \frac{8}{8} = 1 \end{aligned}$$

4. F-Measure

Untuk menghitung F-Measure ini dilakukan dengan menggunakan persamaan (12) berikut.

$$\begin{aligned} f - \text{measure} &= \frac{(2 \times \text{recall} \times \text{presisi})}{(\text{recall} + \text{presisi})} & (13) \\ &= \frac{2 \times 1 \times 0,73}{1 + 0,73} = \frac{1,46}{1,73} = 0,84 \end{aligned}$$

5. KESIMPULAN

Pada penelitian ini penulis telah melakukan pemodelan menggunakan metode *Naïve Bayes classifier*, maka berdasarkan hasil yang diketahui bahwa algoritma *Naïve Bayes* mampu mengklasifikasikan sentimen pada *tweet* dari Divisi Humas Polri. Dengan menggunakan algoritma ini penulis mendapatkan nilai akurasi yang cukup tinggi. Dengan memanfaatkan 122 data *tweet* penulis mendapatkan hasil 61% data positif dan 39% data negatif dengan akurasi sebesar 76%. hasil menunjukkan tingkat nilai positif lebih besar dari nilai negatif. Dengan demikian banyak masyarakat yang merasa puas dengan pelayanan polisi di Indonesia berdasarkan data dari twitter. kemudian peneliti melakukan pengujian terhadap 5 data baru opini masyarakat tentang pelayanan kepolisian dengan memanfaatkan implementasi pada pemodelan, penulis mendapatkan hasil dengan 2 data negatif dan 3 data positif yang kemudian peneliti berharap ini akan menjadi bahan evaluasi bagi pihak kepolisian untuk meningkatkan pelayanan terhadap masyarakat.

Berdasarkan kesimpulan yang telah dituangkan penulis, maka ada beberapa saran yang perlu untuk disampaikan, ini merupakan saran yang diberikan peneliti untuk tindak lanjut atau pengembangan yang lebih mendalam bagi peneliti-peneliti selanjutnya, beriku ini adalah saran untuk mencapai tujuan yaitu :

- 1) Pada penelitian selanjutnya, untuk dapat mempertimbangkan sampel data *tweet* atau komentar pada media sosial untuk dilakukan analisis.

- 2) Pada penelitian selanjutnya, menambahkan model *deployment* atau membangun *website* agar hasil prediksi dapat diakses secara *online*.
- 3) Pada penelitian selanjutnya, disarankan untuk menggunakan lebih dari satu metode untuk melakukan klasifikasi.
- 4) Pada penelitian selanjutnya, proses implementasi dapat memproses data dengan jumlah yang banyak.
- 5) Pada penelitian selanjutnya, menambahkan fitur untuk tweet dari twitter secara *real time* pada model yang akan dibangun.
- 6) Pada Penelitian Selanjutnya, Menambahkan Preprocessing Tokenizing untuk memisahkan kata yang tidak menggunakan *space*.

DAFTAR PUSTAKA

- [1] G. Tika, A. Adiwijaya, and N. D. W. Cahyani, "Klasifikasi Topik Berita Berbahasa Indonesia Menggunakan Multilayer Perceptron," *eProceedings Eng.*, vol. 6, no. 1, pp. 3631–3640, 2019, [Online]. Available: <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/8503>
- [2] Consejería de Educación. Junta de Castilla y León, "濟無No Title No Title No Title," *Protoc. prevención y Organ. del regreso a la Act. lectiva en los centros Educ. Castilla y León para el curso Esc. 2020/2021*, no. 2, pp. 1–19, 2020.
- [3] S. Fransiska, "ANALISIS SENTIMEN TWITTER UNTUK REVIEW FILM MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER (NBC) PADA SENTIMEN R PROGRAMMING," *J. Siliwangi*, vol. 5, no. 2, 2019.
- [4] B. M. Pintoko and K. M. L., "Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naive Bayes Classifier," *e-Proceeding Eng.*, vol. 5, no. 3, pp. 8121–8130, 2018.
- [5] A. Zaki, M. Faiz, "Survei Indikator: Tingkat Kepercayaan Publik terhadap Polri Menurun," *nasional.tempo.co*, 2022. <https://nasional.tempo.co/read/1626759/survei-indikator-tingkat-kepercayaan-publik-terhadap-polri-menurun> (accessed Mar. 28, 2023).
- [6] F. Fathonah and A. Herliana, "Penerapan Text Mining Analisis Sentimen Mengenai Vaksin Covid - 19 Menggunakan Metode Naïve Bayes," *J. Sains dan Inform.*, vol. 7, no. 2, pp. 155–164, Dec. 2021, doi: 10.34128/jsi.v7i2.331.
- [7] O. Penelitian, D. Metode Penelitian, and O. Penelitan, "BAB III."
- [8] S. Nurul, J. Fitriyyah, N. Safriadi, E. Esyudha, and P. #3, "JEPIN (Jurnal Edukasi dan Penelitian Informatika) Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes", [Online]. Available: <http://dev.twitter.com>.
- [9] K. Sihotang and R. Ghaniy, "Penerapan Metode Naïve Bayes Classifier Untuk Penentuan Topik Tugas Akhir Pada Website Perpustakaan STIKOM Binaniaga," *J. Ilm. Teknol. dan Sains*, vol. 9, pp. 63–72, 2019.
- [10] R. Yasmin, "Covid-19 Menggunakan Metode Naive Bayes Classifier Pada Media Sosial Twitter Covid-19 Menggunakan Metode Naive Bayes," 2021.
- [11] M. Imron, "Penerapan Data Mining Algoritma Naives Bayes Dan Part Untuk Mengetahui Minat Baca Mahasiswa Di Perpustakaan Stmik Amikom Purwokerto," *J. Telemat.*, vol. 10, no. 2, pp. 121–135, 2017.
- [12] V. I. Santoso, G. Virginia, and Y. Lukito, "Penerapan Sentiment Analysis Pada Hasil Evaluasi Dosen Dengan Metode Support Vector Machine," *J. Transform.*, vol. 14, no. 2, p. 72, 2017, doi: 10.26623/transformatika.v14i2.439.