

Analisis Sentimen Review Hotel Favehotel Kusumanegara Menggunakan Algoritma *Multinomial Naive Bayes*

Dio Farrel Putra Rachmawan¹, Made Hanindia Prami Swari^{2*}, Chrystia Aji Putra³

^{1,3} Program Studi Informatika, Universitas Pembangunan Nasional “Veteran” Jawa Timur

119081010144@student.upnjatim.ac.id

madehanindia.fik@upnjatim.ac.id

ajiputra@upnjatim.ac.id

*Corresponding author email: madehanindia.fik@upnjatim.ac.id

Abstrak— Pariwisata di Indonesia telah mengalami perkembangan yang pesat dalam beberapa tahun terakhir. Jumlah pengunjung, infrastruktur, fasilitas, dan pelayanan semakin meningkat. Hotel merupakan fasilitas yang sangat penting dalam industri pariwisata, karena berperan sebagai tempat menginap bagi para wisatawan. Salah satu bentuk umpan balik yang paling berharga dalam industri perhotelan adalah ulasan online yang diberikan oleh para pelanggan. Dalam bisnis perhotelan, analisis sentimen untuk *review* hotel menjadi semakin penting dalam rangka memahami keinginan dan kebutuhan konsumen. Analisis sentimen adalah suatu proses untuk mengidentifikasi, mengumpulkan, dan menganalisis opini yang terdapat dalam teks. Dalam beberapa tahun terakhir, metode *Multinomial Naive Bayes* telah menjadi salah satu alat yang populer dan efektif untuk melakukan analisis sentimen. Beberapa keunggulan yang dimiliki oleh *Multinomial Naive Bayes* yaitu kesederhanaan, kecepatan, dan akurasi tinggi, yang memungkinkannya untuk memproses teks dengan baik. Penelitian ini menggunakan dataset mengenai *review* pelanggan hotel yang diambil melalui internet dengan mekanisme *web scrapping* sebanyak 3888 data. Kemudian data tersebut akan melewati tahap *preprocessing* data yang terdiri dari tahap *Cleaning*, *Case Folding*, *Tokenizing*, *Stopword Removal*, dan *Stemming*. Setelah itu dilakukan pembobotan data dengan menggunakan *Term Frequency-Inverse Document Frequency*. Berdasarkan penelitian didapatkan nilai akurasi sebesar 88%. Selain itu, diperoleh pula nilai rata-rata *Precision* sebesar 88%, nilai rata-rata *Recall* sebesar 88%, dan nilai rata-rata *F1-Score* sebesar 87%.

Kata Kunci— Analisis Sentimen, TF-IDF, Text Mining, Naive Bayes Classifier, Multinomial Naive Bayes

I. PENDAHULUAN

Pariwisata di Indonesia telah mengalami perkembangan yang pesat dalam beberapa tahun terakhir. Jumlah pengunjung, infrastruktur, fasilitas, dan pelayanan semakin meningkat. Di seluruh pulau Indonesia terdapat berbagai destinasi wisata menarik dengan keunikan yang khas. Hal ini berdampak pada persaingan bisnis di kota-kota wisata khususnya persaingan dalam bisnis perhotelan. Hotel merupakan fasilitas yang sangat penting dalam industri pariwisata, karena berperan sebagai tempat menginap bagi para wisatawan. Faktor-faktor seperti

fasilitas, kenyamanan, harga, dan ulasan pelanggan sebelumnya mempengaruhi keputusan saat memesan hotel.

Salah satu bentuk umpan balik yang paling berharga dalam industri perhotelan adalah ulasan online yang diberikan oleh para pelanggan. Ulasan ini memberikan pandangan subjektif mengenai pengalaman mereka selama menginap di hotel dan dapat mempengaruhi persepsi calon tamu lainnya. Menurut studi yang dilakukan oleh Chen et al. (2020), disebutkan bahwa ulasan online memiliki dampak signifikan terhadap persepsi dan keputusan pembelian seorang konsumen [1].

Semakin banyak ulasan yang diberikan maka semakin banyak pula waktu yang dibutuhkan untuk menganalisis ulasan secara keseluruhan. Hal ini yang menjadikan proses analisis sentimen dibutuhkan dalam proses menganalisis ulasan yang didapatkan oleh perusahaan ke dalam masing-masing kelas. Dalam bisnis perhotelan, analisis sentimen untuk *review* hotel menjadi semakin penting dalam rangka memahami keinginan dan kebutuhan konsumen. Analisis sentimen adalah suatu proses untuk mengidentifikasi, mengumpulkan, dan menganalisis opini yang terdapat dalam teks. Tugas pokok dalam analisis sentimen adalah mengklasifikasikan polaritas teks dalam dokumen, kalimat, atau pendapat. Polaritas ini mengacu pada apakah teks yang terdapat dalam dokumen, kalimat, atau pendapat memiliki sifat positif atau negatif [2].

Dalam beberapa tahun terakhir, metode *Multinomial Naive Bayes* telah menjadi salah satu alat yang populer dan efektif untuk melakukan analisis sentimen. Metode ini memanfaatkan teori probabilitas untuk mengklasifikasikan teks berdasarkan sentimen yang terkandung di dalamnya. Beberapa keunggulan yang dimiliki oleh *Multinomial Naive Bayes* yaitu kesederhanaan, kecepatan, dan akurasi tinggi, yang memungkinkannya untuk memproses teks dengan baik. Metode MNBC (*Multinomial Naive Bayes Classifier*) digunakan untuk mengklasifikasikan atau mengategorikan teks dengan menggunakan atribut kata yang terdapat dalam dokumen sebagai dasar klasifikasi [3].

Terdapat beberapa penelitian terdahulu mengenai analisis sentimen tentang ulasan pelanggan, seperti halnya penelitian yang dilakukan oleh Sipayung et al., (2016) tentang analisis sentimen komentar pelanggan dengan NBC menghasilkan nilai akurasi sentimen sebesar 77,14% dan 75,42% [4]. Kemudian dalam penelitian lain yang dilakukan oleh Saraswati dan

Riminarsih tentang analisis sentimen terhadap pelayanan KRL *Commuterline* menggunakan algoritma *Bernoulli Naive Bayes* menghasilkan akurasi sebesar 85% [5].

Berdasarkan beberapa hal di atas maka dibuatlah penelitian tentang analisis sentimen *review* pelanggan hotel dengan menggunakan algoritma *Multinomial Naive Bayes*. Tujuan dari penelitian ini yaitu melakukan analisis sentimen terhadap *review* pengunjung hotel dengan menggunakan algoritma *Multinomial Naive Bayes*. Algoritma *Multinomial Naive Bayes* diharapkan dapat memberikan akurasi yang tinggi guna mendapatkan hasil klasifikasi sentimen tepat dan baik.

II. KAJIAN PUSTAKA

A. Penelitian Terdahulu

Sebelum dilakukan penelitian, perlu dilakukan survei dan pembelajaran berdasarkan penelitian terdahulu. Penelitian terdahulu yang pertama yaitu penelitian yang dilakukan oleh Sipayung et al., pada tahun 2016 tentang analisis sentimen komentar pelanggan menggunakan metode *Naive Bayes Classifier*. Dalam penelitian tersebut digunakan dataset sebanyak 175 data komentar yang diambil dari situs web *online*. Hasil akurasi dari klasifikasi pada penelitian tersebut yaitu sebesar 77,14% untuk akurasi kategori dan 75,42% untuk akurasi sentimen [4]. Kemudian penelitian selanjutnya yaitu penelitian yang dilakukan oleh Saraswati dan Riminarsih pada tahun 2020 tentang analisis sentimen terhadap pelayanan KRL *Commuterline* menggunakan algoritma *Bernoulli Naive Bayes*. Penelitian tersebut menggunakan 8130 data yang diambil dari media sosial *twitter* dan didapatkan hasil akurasi sebesar 85% dengan nilai rata-rata *recall* sebesar 85%, nilai rata-rata *precision* sebesar 88,5%, dan nilai *F1-score* sebesar 84,5%.

B. Text Mining

Text mining merupakan area baru dalam pengolahan data yang berupaya untuk mengambil informasi yang signifikan dari teks. Secara konseptual, text mining melibatkan analisis teks guna mengekstrak informasi yang memiliki nilai atau relevansi untuk tujuan tertentu. Text mining mirip dengan data mining, tetapi text mining juga dapat bekerja dengan kumpulan data yang tidak terstruktur atau semi-terstruktur seperti email, dokumen teks, dan file HTML, dll. Oleh karena itu, text mining adalah solusi yang jauh lebih baik [6].

C. Analisis Sentimen

Analisis sentimen adalah kombinasi antara data mining dan text mining, yang digunakan untuk mengolah berbagai pendapat berupa teks yang diberikan oleh konsumen atau pakar melalui berbagai media mengenai produk, layanan, atau instansi tertentu. Analisis sentimen mengidentifikasi tiga jenis pendapat, yaitu pendapat positif, pendapat negatif, dan pendapat netral. Dengan menggunakan analisis sentimen, perusahaan atau instansi terkait dapat memahami tanggapan masyarakat terhadap layanan atau produk tertentu melalui umpan balik yang diberikan oleh masyarakat atau para ahli [4].

Pada analisis sentimen, fokus topik menjadi penting karena pernyataan yang sama dalam topik yang berbeda dapat memiliki makna yang berbeda pula. Oleh karena itu, dalam beberapa penelitian, terutama pada tinjauan produk, langkah

awal dilakukan dengan mengidentifikasi elemen-elemen produk yang dibahas sebelum memulai proses analisis sentimen [7].

D. Preprocessing Data

Tahap *preprocessing* data merupakan tahapan yang dilakukan sebelum masuk ke dalam proses klasifikasi data dengan tujuan untuk mengubah data mentah ke dalam bentuk yang terstruktur atau dapat diproses [5]. Berikut ini adalah beberapa tahapan yang ada dalam *preprocessing* data:

1) Cleaning

Proses ini bertujuan untuk membersihkan atribut data yang tidak memiliki pengaruh untuk tahap klasifikasi seperti tanda baca, karakter kosong, dan emoji [8].

2) Case Folding

Proses ini bertujuan untuk mengonversi semua huruf yang terdapat dalam dokumen menjadi huruf kecil [8].

3) Tokenizing

Proses ini bertujuan untuk menguraikan kalimat menjadi bagian-bagian kata yang terpisah guna memudahkan proses klasifikasi data [8].

4) Stopword Removal

Proses ini bertujuan untuk menghapus kata-kata yang dianggap tidak memiliki relevansi terhadap makna kalimat seperti kata henti [8].

5) Stemming

Proses ini bertujuan untuk mengkonversi kata-kata yang ada dalam dataset menjadi bentuk kata dasar [8]. Proses ini menghilangkan kata imbuhan yang ada dalam dataset.

E. Term Frequency - Inverse Document Frequency

Setiap kata atau fitur yang ada dalam dataset sebelum dilakukan tahap klasifikasi perlu dilakukan pembobotan data. Hal ini bertujuan untuk mengevaluasi tingkat pentingnya kata-kata tersebut dalam konteks analisis sentimen. Salah satu metode pembobotan data yaitu *Term Frequency-Inverse Document Frequency* (TF-IDF).

$$TF - IDF_{t,f} = tf_{t,d} \times \log \frac{N}{df_t} \quad (1)$$

Term Frequency-Inverse Document Frequency (TF-IDF) adalah metode pembobotan yang menggabungkan konsep Term Frequency dan Document Frequency. Konsep Term Frequency digunakan untuk mengukur seberapa sering suatu term muncul dalam satu dokumen dengan memperhitungkan frekuensi kemunculannya. Sementara itu, konsep Document Frequency menghitung jumlah dokumen di mana suatu term muncul. Semakin jarang suatu term muncul dalam dokumen, semakin rendah nilai bobotnya. Dengan menggunakan metode TF-IDF, term dapat diberi bobot yang lebih tinggi jika sering muncul dalam satu dokumen tetapi jarang muncul secara umum dalam koleksi dokumen [9].

F. Multinomial Naive Bayes

Pengklasifikasi Naive Bayes adalah metode klasifikasi yang paling sederhana dan paling banyak digunakan. Model klasifikasi Naive Bayes mengestimasi probabilitas kelas yang relevan berdasarkan distribusi kata-kata dalam dokumen [9]. Berikut adalah bentuk persamaan *Naive Bayes* secara umum:

$$P(H|X) = \frac{P(H)P(X|H)}{P(X)} \quad (2)$$

Salah satu tipe algoritma *Naive Bayes Classifier* yaitu *Multinomial Naive Bayes*. Algoritma *Multinomial Naive Bayes* merupakan algoritma yang sering digunakan untuk menyelesaikan pengklasifikasian teks ataupun dokumen. Metode ini memanfaatkan teori probabilitas untuk mengklasifikasikan teks berdasarkan sentimen yang terkandung di dalamnya. Beberapa keunggulan yang dimiliki oleh *Multinomial Naive Bayes* yaitu kesederhanaan, kecepatan, dan akurasi tinggi, yang memungkinkannya untuk memproses teks dengan baik. Metode MNBC (*Multinomial Naive Bayes Classifier*) digunakan untuk mengklasifikasikan atau mengategorikan teks dengan menggunakan atribut kata yang terdapat dalam dokumen sebagai dasar klasifikasi [3].

G. Confusion Matrix

Confusion Matrix merupakan salah satu metode yang digunakan untuk evaluasi model dalam pemrosesan data yang memiliki keluaran dengan dua kelas atau lebih.

TABEL I
CONFUSION MATRIX

		Predicted Values	
		Predicted Negative	Predicted Positive
Actual Values	Actual Negative	True Negative	False Positive
	Actual Positive	False Negative	True Positive

Berdasarkan Tabel I *confusion matrix* memiliki 4 hasil penilaian yaitu *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN). Hal itu didapatkan berdasarkan perbandingan nilai prediksi dan nilai aktual.

Kemudian seluruh hasil tersebut digunakan untuk menghitung nilai performansi seperti:

1) Accuracy

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (3)$$

2) Precision

$$Precision = \frac{TP}{(TP + FP)} \quad (4)$$

3) Recall

$$Recall = \frac{TP}{(TP + FN)} \quad (5)$$

4) F1-Score

$$F - 1 \text{ Score} = \frac{(2 \times Recall \times Precision)}{(Recall + Precision)} \quad (6)$$

III. METODOLOGI PENELITIAN

A. Pengumpulan Dataset

Penelitian ini menggunakan dataset mengenai *review* pelanggan hotel yang diambil melalui internet dengan mekanisme *web scrapping*. Data diambil melalui website www.traveloka.com dari salah satu hotel yang ada di Indonesia yaitu Favehotel Kusumanegara Yogyakarta. Favehotel Kusumanegara Yogyakarta adalah sebuah hotel bintang 3 yang terletak di Jl. Kusumanegara 91 Umbulharjo Yogyakarta, pusat kota Yogyakarta. Hotel ini memiliki lokasi yang strategis di dekat berbagai destinasi wisata populer seperti Malioboro, Kraton Yogyakarta, Kebun Binatang Gembira Loka, dan lainnya. Jumlah data yang berhasil diambil yaitu sejumlah 3888 data. Selanjutnya data diberikan label agar komputer dapat dengan mudah melatih data.

B. Preprocessing Data

Setelah seluruh data diambil dan diberikan label, selanjutnya data melewati tahap *preprocessing* sebelum nantinya data diklasifikasi.

Contoh hasil setelah melewati tahap *preprocessing* data dapat dilihat pada Tabel II.

TABEL II
CONTOH HASIL PREPROCESSING DATA

Proses	Hasil
Cleaning	Tempat bersih nyaman pelayanan bagus Sebagian besar lorong perlu dirapikan lagi Terima kasih
Case Folding	tempat bersih nyaman pelayanan bagus sebagian besar lorong perlu dirapikan lagi terima kasih
Tokenizing	['tempat', 'bersih', 'nyaman', 'pelayanan', 'bagus', 'sebagian', 'besar', 'lorong', 'perlu', 'dirapikan', 'lagi', 'terima', 'kasih']
Stopword Removal	['tempat', 'bersih', 'nyaman', 'pelayanan', 'bagus', 'sebagian', 'besar', 'lorong', 'dirapikan', 'terima', 'kasih']
Stemming	['tempat', 'bersih', 'nyaman', 'layan', 'bagus', 'bagi', 'besar', 'lorong', 'rapi', 'terima', 'kasih']

C. Pembobotan Data

Setelah melewati tahap *preprocessing*, selanjutnya akan dilakukan proses pembobotan data dengan menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Dalam proses ini, untuk mendapatkan nilai bobot w perlu diketahui terlebih dahulu nilai *term frequency* yang merupakan bobot *term* t dalam dokumen d . Kemudian dilakukan pencarian nilai *idf* yang merupakan banyak dokumen yang berisi istilah yang dicari. Setelah kedua nilai didapatkan, selanjutnya kedua nilai tersebut dikalikan dan didapatkanlah hasil nilai bobot w .

```
# Mengimport CountVectorizer dan TfidfVectorizer dari library sklearn
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfVectorizer

# Proses TF-IDF
data_clean = data
data_clean = data_clean.astype({'Sentiment' : 'category'})
data_clean = data_clean.astype({'Ulasan' : 'string'})

tf = TfidfVectorizer()
text_tf = tf.fit_transform(data_clean['Ulasan'].astype('U'))
tokens_tf = tf.get_feature_names_out()
data = pd.DataFrame(data = text_tf.toarray(), columns = tokens_tf)

data
```

Gbr. 1 Potongan Kode Program TF-IDF

D. Klasifikasi Data

Sebelum data masuk ke dalam proses klasifikasi, dataset yang telah disiapkan dibagi terlebih dahulu ke dalam dua subset data yaitu data latih dan data uji. Perbandingan dalam pembagian data latih dan data uji pada penelitian ini yaitu sebesar 90% dan 10%. Kemudian, setelah seluruh data siap maka selanjutnya dilakukan proses klasifikasi data dengan menggunakan algoritma *Multinomial Naive Bayes*. Pada tahap ini nantinya akan dibuatkan sebuah *class* dimana dalam *class* tersebut terdapat fungsi pelatihan dan juga fungsi prediksi. Pada proses klasifikasi digunakan nilai alpha untuk mencegah terjadinya nol mutlak atau probabilitas yang sangat rendah dalam perhitungan probabilitas posterior. Nilai alpha yang digunakan dalam klasifikasi ini yaitu 1,0.

IV. HASIL DAN PEMBAHASAN

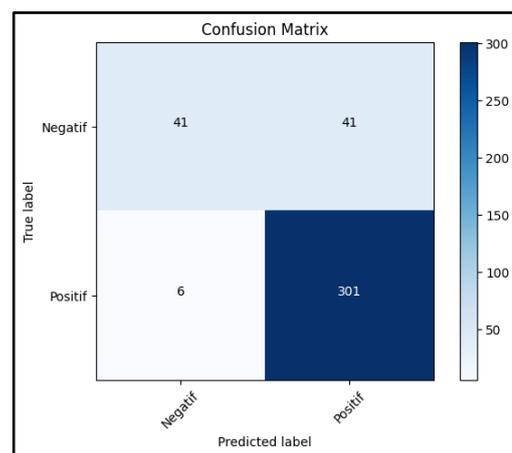
Proses penelitian dimulai dari pengumpulan data dengan metode *web scrapping* pada situs web www.traveloka.com untuk mendapatkan data ulasan salah satu hotel yang ada di Indonesia. Jumlah data yang siap diproses yaitu sebanyak 3888 data yang telah memiliki label positif dan negatif. Kemudian dilakukan *preprocessing* seluruh data meliputi proses *cleaning*, *case folding*, *tokenizing*, *stopword removal*, dan *stemming* yang dapat dilihat pada Tabel II.

Setelah seluruh data selesai melewati tahap *preprocessing* selanjutnya dilakukan pembobotan data dengan TF-IDF dan didapatkan hasil proses TF-IDF seperti Gbr.2

	abai	abuabu	ac	acak	acara	accura	acnya	acung	ada	adem	...	yah	yaw	yg	yoga	yo
0	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.328247	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0
...
3882	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0
3883	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0
3884	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0
3885	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0
3886	0.0	0.0	0.000000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0

Gbr. 2 Hasil Proses TF-IDF

Kemudian seluruh data dibagi ke dalam 2 subset yaitu data latih dan data uji dengan perbandingan sebesar 90% untuk data latih dan 10% untuk data uji. Setelah dibagi seluruh data diklasifikasi dengan menggunakan algoritma *Multinomial Naive Bayes* serta dilakukan evaluasi model dengan menggunakan *confusion matrix*. Berdasarkan hasil evaluasi model didapatkan nilai akurasi klasifikasi data sebesar 88% seperti yang ditunjukkan pada Gbr. 3



Gbr. 3 Hasil *Confusion Matrix*

Pada hasil klasifikasi, model berhasil mengklasifikasikan dengan tepat 342 dari total 389 data. Dalam dataset tersebut, terdapat 301 data dengan label positif dan 41 data dengan label negatif. Nilai *Precision* yang diperoleh adalah 87% untuk sentimen negatif dan 88% untuk sentimen positif. Ini berarti dari 47 data yang diprediksi sebagai sentimen negatif, 41 data berhasil diprediksi dengan benar sebagai negatif, dan 6 data diprediksi secara salah. Selanjutnya, dari 342 data yang diprediksi sebagai sentimen positif, 301 data berhasil diprediksi dengan benar sebagai positif, dan 41 data diprediksi secara salah.

Classification Report :				
	precision	recall	f1-score	support
0	0.87	0.50	0.64	82
1	0.88	0.98	0.93	307
accuracy			0.88	389
macro avg	0.88	0.74	0.78	389
weighted avg	0.88	0.88	0.87	389

Gbr. 4 Hasil *Classification Report*

Untuk nilai *Recall*, diperoleh hasil sebesar 50% untuk sentimen negatif dan 93% untuk sentimen positif. Ini berarti dari 82 data yang seharusnya memiliki sentimen negatif, model berhasil mengidentifikasi 41 data dengan benar, tetapi tidak dapat mengidentifikasi 41 data lainnya. Selanjutnya, dari 307 data yang seharusnya memiliki sentimen positif, model berhasil mengidentifikasi 301 data dengan benar, tetapi tidak dapat mengidentifikasi 6 data lainnya.

Setelah mendapatkan nilai Precision dan Recall, juga diperoleh nilai F1-Score yang merupakan nilai rata-rata harmonik dari precision dan recall. Untuk skenario 1, diperoleh F1-Score sebesar 64% untuk sentimen negatif dan 93% untuk sentimen positif.

V. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan maka dapat ditarik kesimpulan bahwa berdasarkan 3888 data *review* salah satu hotel di Indonesia dan telah dilakukan tahap klasifikasi data menggunakan *Multinomial Naive Bayes* didapatkan nilai akurasi sebesar 88%. Selain itu, diperoleh pula nilai rata-rata *Precision* sebesar 88%, nilai rata-rata *Recall* sebesar 88%, dan nilai rata-rata *F1-Score* sebesar 87%.

Adapun saran untuk penelitian selanjutnya yaitu dataset yang digunakan dapat dikembangkan. Lalu, pada tahap perancangan model dapat ditambahkan metode lain sebagai perbandingan untuk mengetahui hasil akurasi terbaik dari beragam metode. Selain itu, model ini dapat diperluas dan dikembangkan lebih lanjut untuk diterapkan pada berbagai aplikasi atau media lainnya, sehingga tidak terbatas hanya pada tahap penelitian ini.

UCAPAN TERIMA KASIH

Penulis ingin mengucapkan rasa terima kasih yang sebesar-besarnya kepada Rektor UPN "Veteran" Jawa Timur, para Dosen Fakultas Ilmu Komputer, dan Koordinator Program Studi Informatika atas dukungan yang diberikan dalam

penelitian ini dan juga kesempatan yang diberikan untuk melakukan penelitian ini hingga dapat menghasilkan jurnal yang relevan.

REFERENSI

- [1] Chen, M.-Y., Teng, C.-I., & Chiou, K.-W. (2020). The helpfulness of online reviews Images in review content and the facial expressions of reviewers' avatars. *Online Information Review*, 44(1), 90-113. <https://doi.org/10.1108/OIR-08-2018-0251>
- [2] Nugroho, D. G., Chrisnanto, Y. H., & Wahana, A. (2016). ANALISIS SENTIMEN PADA JASA OJEK ONLINE MENGGUNAKAN METODE NAÏVE BAYES. *Seminar Nasional Sains dan Teknologi*, 1(1), 156-161. <http://dx.doi.org/10.36499/psnst.v1i1.1526>
- [3] Mandasari, S., Hayadi, B. H., & Gunawan, R. (2022). Analisis Sentimen Pengguna Transportasi Online Terhadap Layanan Grab Indonesia Menggunakan Multinomial Naive Bayes Classifier. *Jurnal Teknologi Sistem Informasi dan Sistem Komputer TGD*, 5(2), 118-126. <https://doi.org/10.53513/jsk.v5i2.5635>
- [4] Sipayung, E. M., Maharani, H., & Zefanya, I. (2016). PERANCANGAN SISTEM ANALISIS SENTIMEN KOMENTAR PELANGGAN MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER. *Jurnal Sistem Informasi (JSI)*, 8(1), 985-965. <https://doi.org/10.36706/jsi.v8i1.3250>
- [5] Saraswati, M., & Rimirasih, D. (2020). ANALISIS SENTIMEN TERHADAP PELAYANAN KRL COMMUTERLINE BERDASARKAN DATA TWITTER MENGGUNAKAN ALGORITMA BERNOULLI NAIVE BAYES. *Jurnal Ilmiah Informatika Komputer*, 25(3), 225-238. <http://dx.doi.org/10.35760/ik.2020.v25i3.3256>
- [6] Kumar, L. (2013). TEXT MINING: CONCEPTS, PROCESS AND APPLICATIONS. *Journal of Global Research in Computer Science*, 4(3), 36-39.
- [7] *Bayesian Opinion Mining – Accidental Factors – Needless text.* (2010, January 21). Accidental Factors. Retrieved June 22, 2023, from <https://accidentalfactors.com/bayesian-opinion-mining/>
- [8] Hendriyanto, M. D., Ridha, A. A., & Enri, U. (2022). ANALISIS SENTIMEN ULASAN APLIKASI MOLA PADA GOOGLE PLAY STORE MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE. *INTECOMS: Journal of Information Technology and Computer Science*, 5(1), 1-7. <https://doi.org/10.31539/intecom.v5i1.3708>
- [9] Yulita, W., Nugroho, E. D., & Alghifari, M. H. (2021). Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier. *JDMSI: Jurnal Data Mining dan Sistem Informasi*, 2(2), 1-9. <https://doi.org/10.33365/jdmsi.v2i2.1344>