

Paper JSON

by Sunneng Sandino

Submission date: 07-Jan-2023 01:23PM (UTC+0700)

Submission ID: 1989445387

File name: Paper_json.pdf (1.02M)

Word count: 4583

Character count: 26918

Text Mining dan Klasifikasi Sentimen Berbasis Naïve Bayes Pada Opini Masyarakat terhadap Makanan Tradisional

Sunneng Sandino Berutu

Fakultas Sains dan Komputer, Program Studi Informatika, Universitas Kristen Immanuel, Yogyakarta, Indonesia

Email: sandinoberutu@ukrimuniversity.ac.id

Email Penulis Korespondensi: sandinoberutu@ukrimuniversity.ac.id

Submitted: 17/11/2022; Accepted: 12/12/2022; Published: 31/12/2022

Abstrak—Indonesia memiliki berbagai jenis makanan tradisional terkenal dan tersedia di berbagai kota. Selain makanan khas tersebut, beberapa jenis makanan internasional juga sudah merambah ke seluruh pelosok nusantara. Untuk menganalisis sentimen masyarakat terhadap kedua kategori makanan tersebut, peneliti melakukan *text mining* dengan mengambil data dari twitter. Jenis makanan tradisional yang dianalisis adalah rendang, sate dan gudeg. Sementara itu, jenis makanan internasional terdiri dari pizza, hamburger dan spaghetti. Tahapan yang dilakukan dalam penelitian ini meliputi *data crawling*, *data cleaning*, *data filtering*, *data translating*, *data splitting* dan pengembangan model klasifikasi teks berbasis Naïve Bayes. Hasil analisis menunjukkan persentase sentimen positif tertinggi terdapat pada makanan gudeg sebesar 57,9. Kemudian, persentase sentimen negatif tertinggi diperoleh makanan rendang sebesar 21,9. Selanjutnya, persentase tertinggi sentimen netral dimiliki oleh makanan hamburger. Sementara itu, kinerja model klasifikasi dengan dataset hamburger memperoleh nilai tertinggi pada parameter seperti *accuracy* sebesar 0,72, *precision* sebesar 0,72 dan *recall* sebesar 0,68 dalam memprediksi data sentimen.

Kata Kunci: Makanan Tradisional; Makanan Internasional; Sentimen; Text Mining; Twitter; Naïve Bayes

Abstract—Indonesia has several famous traditional foods and is available in some cities. In addition, several international foods also are interesting to Indonesian. This article analyzes the netizen sentiment for these food categories where the data source is Twitter. The foods are rendang, sate, gudeg, pizza, hamburger, and spaghetti. The text mining approach is adopted to process data. The research steps are data crawling, cleaning, filtering, translating, and splitting. Furthermore, the classifier model based on the Naïve Bayes algorithm is developed. The analysis result shows that the gudeg food reaches a high percentage of positive sentiment with 57.9. Then, the high rate of negative sentiment is achieved by the rendang food with 21.9%. Moreover, hamburger food obtains a high percentage of neutral sentiment. Meanwhile, the evaluation of classifier model performance shows that the model with the hamburger dataset achieves a high score for accuracy, precision, and recall parameters with 0.72, 0.72, and 0.68 sequentially.

Keywords: Traditional Food; International Food; Sentiment; Twitter; Naïve Bayes

1. PENDAHULUAN

Data yang beredar di media sosial sangat besar. Data tersebut dapat diolah dan dimanfaatkan untuk memperoleh berbagai informasi. Salah satu cara untuk mengolah data tersebut dengan menerapkan teknologi kecerdasan buatan yaitu pendekatan *text mining* [1]. Pendekatan ini menjadi salah satu bagian penting dalam penelitian terkait dengan analisis sentimen. Beberapa penelitian terdahulu telah menerapkan *text mining* pada data media sosial seperti respon pengguna terhadap munculnya teroris ISIS [2], sentimen penduduk terhadap kebijakan kenaikan harga dasar listrik [3], dan mengukur keraguan masyarakat terhadap vaksin covid-19 [4].

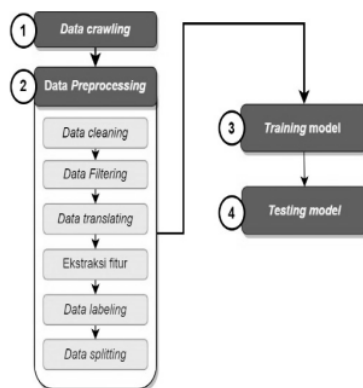
Analisis sentimen pada media sosial menggali opini, perasaan dan sikap pengguna untuk mendapatkan pengetahuan atau informasi berguna dari opini seseorang atau komunitas [5]. Informasi tersebut dapat dipergunakan sebagai bagian dari strategi atau pendukung pengambilan keputusan. Analisis sentimen dengan data bersumber dari media sosial seperti twitter, instagram, facebook dan you tube sudah dilakukan oleh peneliti terdahulu. Twitter menjadi salah satu media sosial populer karena menyediakan berbagai macam fasilitas dan memudahkan pemakai dalam penyampaian informasi ke publik [6]. Data dari media sosial twitter sudah diimplementasikan untuk mengukur sentimen dalam berbagai bidang kehidupan manusia. Untuk bidang pendidikan, peneliti menganalisis respon masyarakat terhadap kegiatan belajar mengajar secara *hybrid* [7]. Selanjutnya, analisis sentimen dapat menjadi salah satu bagian evaluasi terhadap kegiatan pembelajaran di perguruan tinggi [8]. Penerapan di bidang kesehatan, seperti pengukuran pandangan masyarakat terhadap pelaksanaan vaksin covid-19 [9], respon masyarakat terhadap penyakit covid-19 [10]. Selanjutnya, studi kasus pada unsur pemerintahan, misalnya: penilaian masyarakat terhadap kinerja BMKG [11], kinerja anggota DPR [12], dan kinerja jajaran di kementerian kesehatan [13]. Selain itu, beberapa peneliti menganalisis pandangan masyarakat terhadap kebijakan-kebijakan yang dikeluarkan pemerintah [14]–[16], bidang pariwisata [17], dan bidang pemasaran [18].

Naïve Bayes merupakan salah satu algoritma populer *machine learning* yang telah diimplementasikan pada masalah klasifikasi. Algoritma ini memiliki cara kerja yang sederhana, menghasilkan probabilitas akurasi klasifikasi sangat tinggi dan mampu mengurangi konsumsi energi pada tahap pelatihan [19]. Algoritma ini telah diimplementasikan pada klasifikasi teks seperti, identifikasi berita atau informasi hoaks di media digital [20], deteksi spam pada email [21]. Selain itu, untuk klasifikasi sentimen dari data di media sosial telah dilakukan beberapa peneliti terdahulu seperti sentimen produk *e-commerce* dari data ulasan konsumen [22], pengukuran sentimen bidang pemasaran online [23] dan pengukuran kinerja polisi [24].

Indonesia memiliki kekayaan beraneka ragam makanan tradisional yang tersebar di seluruh pelosok nusantara. Cita rasa dan proses pengolahan yang unik membuat makanan tersebut menjadi populer baik di dalam maupun di luar negeri, seperti rendang, sate, gudeg, dan sebagainya. Makanan tradisional merupakan salah satu jati diri bangsa yang menjadi pembeda dengan bangsa lain. Sementara itu, makanan internasional dalam bentuk siap saji juga sudah merambah sejak berpuluh-puluh tahun lalu di seluruh kota di Indonesia, seperti pizza, hamburger, spaghetti dan sebagainya. Umumnya, makanan ini sangat digandrungi oleh kaum milenial. Meskipun demikian, makanan tradisional tetap perlu dipertahankan dan ditingkatkan popularitasnya. Penelitian terkait sentimen masyarakat terhadap kuliner Indonesia telah dilakukan sebelumnya, seperti sentimen jajanan yang tersedia di pusat perkotaan Surabaya[25] dan Bandung [26], penentuan sentimen dengan data *hash tag* kuliner[27], sentimen pada tempat penyediaan kuliner [28] dan restoran [29]. Namun, pengukuran sentimen masyarakat terhadap makanan tradisional dan makanan internasional belum pernah dilakukan. Oleh karena itu, penelitian ini menganalisis opini masyarakat terhadap popularitas makanan tradisional dan makanan internasional dengan melakukan *text mining* di mana data bersumber dari twitter. Selanjutnya, algoritma Naïve Bayes diadopsi pada model klasifikasi sentimen.

⁸ 2. METODOLOGI PENELITIAN

Penelitian ini dilakukan dalam empat tahapan, yaitu ⁴ *data crawling*, *data preprocessing*, *training model* dan *testing model* serta evaluasi model. Tahapan penelitian ini ditunjukkan dalam Gambar 1 berikut.



Gambar 1. Tahapan pelaksanaan kegiatan penelitian

2.1 Data Crawling

Untuk memperoleh data dari twitter, modul *library scraper for social networking service* (SNS) python diadopsi, kemudian fungsi *TwitterSearchScraper* dipanggil. Kata kunci yang dipakai yaitu rendang, sate, gudeg, pizza, hamburger, dan spaghetti. Untuk memastikan data yang ditarik berbahasa Indonesia saja maka property *language* diatur menjadi **lang: id**. Dalam proses *crawling*, tiga properti twitter seperti *tweet.date*, *tweet.content* dan *tweet.lang* diunduh.

2.2 Data Preprocessing

Pada tahap ini, proses-proses yang dilakukan, seperti *data cleaning*, *data filtering*, *data translating*, ekstraksi fitur, *data labeling*, dan *data splitting*. *Library* yang dipergunakan pada *data preprocessing* adalah *TextBlob*.

2.2.1 Data Cleaning

Tahapan ini merupakan proses pembersihan data terhadap teks yang berulang. Selanjutnya, menghapus teks yang tidak penting untuk dianalisis seperti *symbol-symbol*, *mentions*, *hash tag*, *hyperlink* dan beberapa tanda baca. Proses lainnya adalah mengubah teks menjadi huruf kecil.

2.2.2 Data Filtering

Sebelum proses *filtering*, kalimat terlebih dahulu dipisahkan ke dalam kata-kata (*tokenizing*). Selanjutnya, kata-kata yang kurang memiliki makna seperti “dan”, “yang”, “dari”, “di” dan sebagainya dihilangkan.

2.2.3 Data Translating

Untuk memenuhi persyaratan dalam penggunaan *Textblob* berbahasa Inggris maka data perlu diterjemahkan terlebih dahulu ke dalam bahasa Inggris. Untuk proses translasi, *library deep translator* diadopsi dalam penelitian ini. Fungsi *GoogleTranslator* dipanggil untuk menterjemahkan teks dari bahasa Indonesia menjadi bahasa Inggris.

2.2.4 Ekstraksi Fitur

Proses ekstraksi fitur akan menghasilkan nilai *subjectivity* dan *polarity* dalam sebuah kalimat. Untuk memperoleh nilai ke dua parameter tersebut maka fungsi **sentiment.subjectivity** dan **sentiment.polarity** textBlob diadopsi. Rentang nilai *polarity* berada antara [-1,1] sedangkan nilai *subjectivity* terletak diantara [0,1].

2.2.5 Data Labeling

Sentimen dikelompokkan ke dalam tiga kategori, yaitu positif, negatif dan netral [30]. Penentuan label didasarkan pada kondisi berikut ini, yaitu:

- jika nilai *polarity* > 0 maka label = positif.
- jika nilai *polarity* < 0 maka label = negatif.
- jika nilai *polarity* = 0 maka label = netral.

2.2.6 Data Splitting

Data pada pelatihan model menggunakan type data seimbang. Untuk memperoleh data seimbang, pertama-tama dicari hasil perolehan jumlah terkecil diantara sentimen positif, negatif dan netral. Selanjutnya, data yang sudah seimbang dibagi ke dalam data *training* dan data *testing*.

2.3 Training Model

Metode klasifikasi Naïve Bayes merupakan algoritma *machine learning* berbasis teorema Bayesian[31]. Teori ini digunakan untuk menemukan probabilitas dari hipotesis berdasarkan fakta yang ada. Persamaan probabilitas (1) dinyatakan sebagai berikut.

$$P(\text{label}|\text{features}) = \frac{P(\text{features}|\text{label}) \times P(\text{label})}{P(\text{features})} \quad (1)$$

Sintax metode di *library textblob* adalah `classifiers.NaiveBayesClassifier`.

2.4 Testing Model

Model dievaluasi dengan menginputkan data baru. Pengukuran kinerja model pada tahap pengujian ditampilkan dengan pendekatan *confusion matrix*[32]. *Confusion matrix* akan menghasilkan beberapa parameter seperti akurasi, presisi, *recall* dan *f1-score*.

3. HASIL DAN PEMBAHASAN

Bagian ini berisi penjelasan tentang hasil implementasi metode penelitian yang sudah dirancang. Selanjutnya, hasil tersebut dianalisis untuk menggali informasi terkait dengan sentimen masyarakat Indonesia terhadap makanan tradisional dan internasional.

3.1 Analisis Sentimen

Untuk proses *crawling*, kata kunci yang digunakan sesuai dengan nama jenis makanannya, seperti “rendang”, “sate”, “gudeg”, “pizza”, “hamburger”, dan “spaghetti”. Hasil proses *crawling* setiap jenis makanan dideskripsikan dalam Tabel 1 berikut.

Tabel 1. Jenis data tweet

No	Jenis Makanan	Tanggal	Jumlah tweet
1	Rendang	27/10/2022 – 5/11/2022	5000
2	Sate	01/11/2022 – 5/11/2022	5000
3	Gudeg	15/09/2022 – 5/11/2022	5000
4	Pizza	29/10/2022 – 5/11/2022	5000
5	Hamburger	27/05/2022 – 5/11/2022	5000
6	Spaghetti	15/09/2022 – 5/11/2022	5000

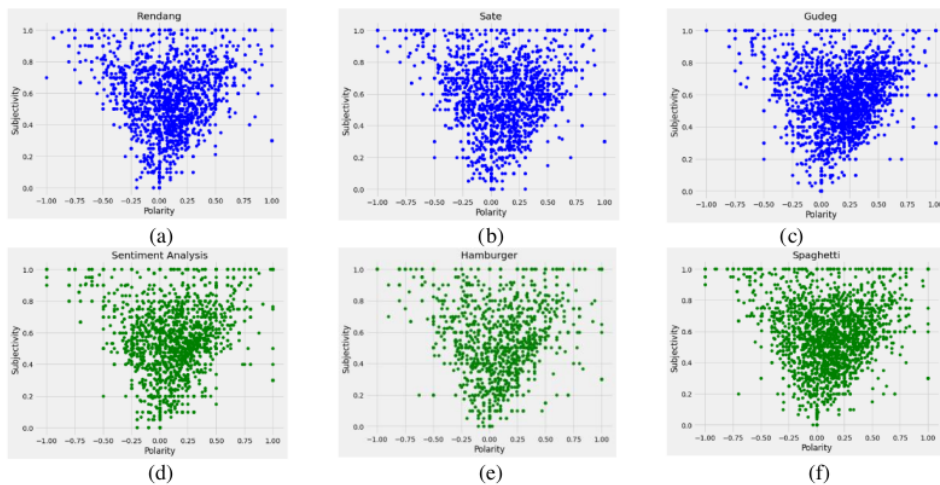
Berdasarkan Tabel 1 di atas, untuk memperoleh tweet sebanyak 5000 data maka jenis makanan sate memiliki rentang waktu terpendek dibandingkan dengan jenis makanan lainnya, sementara itu, makanan hamburger memiliki rentang waktu terlama.

Kata-kata yang ada dalam tweet divisualisasikan dengan menggunakan *library wordcloud* seperti yang terlihat dalam Gambar 2 berikut. Kata kunci setiap jenis makanan memiliki ukuran huruf yang lebih besar diantara kata-kata yang ada. Hal tersebut menunjukkan bahwa kemunculan kata kunci menjadi yang terbanyak dibandingkan dengan kata lainnya.



Gambar 2. Visualisasi *wordcloud* jenis makanan a) rendang, b) sate, c) gudeg, d) pizza, e) hamburger dan f) spaghetti

Untuk menentukan sentimen pada setiap tweet maka nilai *subjectivity* dan *polarity* diukur. Hasil pengukuran kedua variabel tersebut menjadi dasar dalam penentuan label sentimen: "Positif", "Negatif", dan "Netral". Sebaran nilai *subjectivity* dan *polarity* setiap jenis makanan ditampilkan dalam bentuk grafik yang terlihat pada Gambar 3 berikut. Secara umum, sebaran nilai *subjectivity* lebih banyak berada pada rentang 0,4 sampai 1. Hal ini menunjukkan bahwa *tweet* yang diperoleh lebih subjektif. Sementara itu, sebaran nilai *polarity* lebih banyak berada pada rentang antara -0,50 dan 0,50.

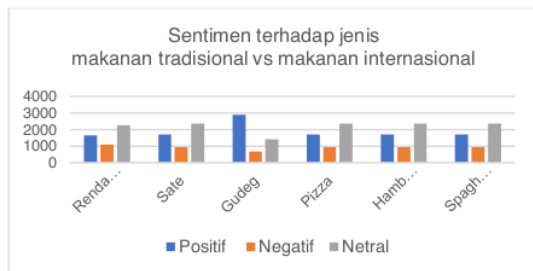


Gambar 3. Korelasi antara nilai *subjectivity* dan *polarity* pada jenis makanan a) rendang, b) sate, c) gudeg, d) pizza, e) hamburger, dan f) spaghetti

Berdasarkan hasil perhitungan nilai *polarity* maka diperoleh hasil analisis sentimen yang tersaji dalam Tabel 2 dan Gambar 4 berikut. Makanan rendang memperoleh persentase sentimen Positif tertinggi sebesar 57,9 sedangkan terendah adalah makanan hamburger. Sementara itu, makanan rendang memiliki persentase sentimen Negatif tertinggi sebesar 21,9 sedangkan sentimen terendah pada makanan gudeg dan hamburger yaitu sebesar 13,7.

Tabel 2. Hasil analisis sentimen

No	Jenis Makanan	Sentimen (%)		
		Positif	Netral	Negatif
1	Rendang	32,9	45,2	21,9
2	Sate	33,9	47,4	18,7
3	Gudeg	57,9	28,3	13,7
4	Pizza	32,3	53,4	14,3
5	Hamburger	25,4	60,9	13,7
6	Spaghetti	40,9	40,6	18,5



Gambar 4. Grafik bar analisis sentimen makanan tradisional dengan makanan internasional

Berdasarkan hasil analisis sentimen, meskipun rentang waktu pengambilan data *tweet* makanan gudeg lebih lama dibanding dengan makanan lainnya, akan tetapi jenis makanan ini memperoleh persentase sentimen positif tertinggi dibanding dengan jenis makanan lainnya. Sementara itu, makanan rendang memperoleh persentase sentimen negatif tertinggi dibandingkan dengan jenis makanan lainnya.

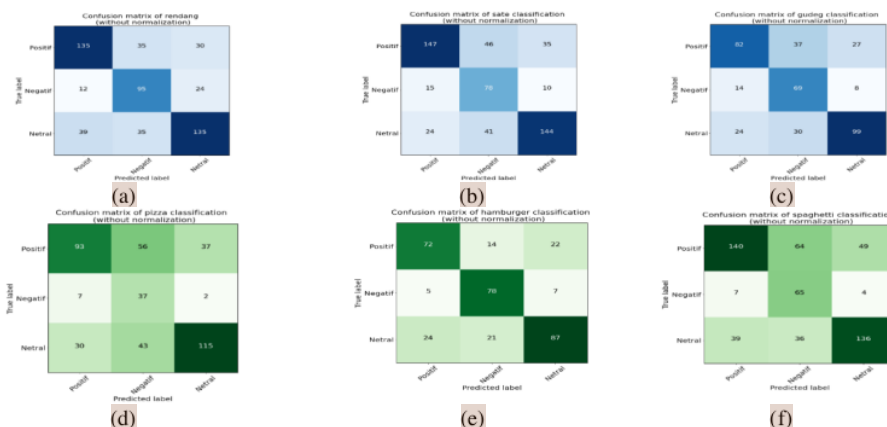
3.2 Model Klasifikasi

Model klasifikasi menggunakan metode Naïve Bayes dengan *dataset balanced*. Untuk memenuhi dataset *balanced* maka jumlah sentimen positif, negatif, dan netral ditentukan berdasarkan jumlah data sentimen terkecil. Jumlah sentimen yang digunakan pada tahap pelatihan dan pengujian model dideskripsikan dalam Tabel 3 berikut. Makanan rendang, sate dan spaghetti memiliki jumlah dataset terbesar yaitu 2700. Dataset dibagi menjadi dua bagian dimana sebesar 80% diperuntukan pada tahap pelatihan sedangkan sisanya digunakan pada tahap pengujian model.

Tabel 3. Deskripsi *dataset balanced* setiap makanan

No	Nama Dataset	Jumlah Data Pelatihan dan Pengujian			Total
		Positif	Netral	Negatif	
1	Rendang	900	900	900	2700
2	Sate	900	900	900	2700
3	Gudeg	650	650	650	1950
4	Pizza	700	700	700	2100
5	Hamburger	650	650	650	1950
6	Spaghetti	900	900	900	2700

Berdasarkan dataset di Tabel 3 maka model akan dilatih dengan 6 dataset yang berbeda sehingga penelitian ini menghasilkan 6 model klasifikasi. Setelah tahapan pelatihan selesai dilakukan, selanjutnya model diujikan dengan dataset pengujian. Hasil pengujian divisualisasikan ke dalam *confusion matrix* seperti yang tersaji dalam Gambar 5 berikut.



Gambar 5. Visualisasi *confusion matrix* jenis makanan a) rendang, b) sate, c) gudeg, d) pizza, e) hamburger dan f) spaghetti

Confusion matrix pada Gambar 5 diatas menjadi acuan dalam menentukan besaran nilai parameter seperti *precision*, *recall* dan *f1-score*. Deskripsi hasil pengukuran parameter pada setiap model tersaji dalam Tabel 4-9.

Hasil prediksi model dengan dataset rendang tersaji dalam Tabel 4 berikut. Sentimen negatif memperoleh nilai *precision* tertinggi sedangkan sentimen positif mencapai nilai tertinggi pada parameter *recall* dan *f1-score*. Sementara itu, nilai rata-rata *accuracy* prediksi sebesar 0,68.

Tabel 4. Hasil pengujian model dengan dataset rendang

No	Nama Label	Parameter		
		Precision	Recall	F1-score
1	Positif	0,68	0,73	0,7
2	Negatif	0,73	0,58	0,64
3	Netral	0,65	0,71	0,68

Deskripsi hasil prediksi model dengan dataset gudeg tersaji dalam Tabel 6 berikut. Nilai *precision* tertinggi diperoleh sentimen negatif sedangkan nilai tertinggi parameter pada *recall* dan *f1-score* dicapai label sentimen netral. Sementara itu, nilai rata-rata akurasi prediksi sebesar 0,68.

Tabel 5. Hasil pengujian model dengan dataset sate

No	Nama Label	Parameter		
		Precision	Recall	F1-score
1	Positif	0,64	0,79	0,71
2	Negatif	0,76	0,47	0,58
3	Netral	0,69	0,76	0,72

Deskripsi hasil prediksi model dengan dataset gudeg tersaji dalam Tabel 6 berikut. Nilai *precision* tertinggi diperoleh sentimen negatif sedangkan nilai tertinggi parameter pada *recall* dan *f1-score* dicapai label sentimen netral. Sementara itu, nilai rata-rata akurasi prediksi sebesar 0,68.

Tabel 6. Hasil pengujian model dengan dataset gudeg

No	Nama Label	Parameter		
		Precision	Recall	F1-score
1	Positif	0,56	0,68	0,62
2	Negatif	0,76	0,51	0,61
3	Netral	0,65	0,74	0,69

Evaluasi kinerja model dengan dataset pizza dideskripsikan dalam Tabel 7 berikut. Nilai tertinggi pada parameter *precision* dicapai sentimen negatif sedangkan nilai tertinggi pada parameter *recall* dan *f1-score* diperoleh sentimen netral. Nilai akurasi rata-rata prediksi sebesar 0,58.

Tabel 7. Hasil pengujian model dengan dataset pizza

No	Nama Label	Parameter		
		Precision	Recall	F1-score
1	Positif	0,5	0,72	0,59
2	Negatif	0,8	0,27	0,41
3	Netral	0,61	0,75	0,67

Deskripsi hasil pengujian model dengan dataset hamburger tersaji dalam Tabel 8 berikut. Nilai tertinggi parameter *precision* dan *f1-score* dicapai sentimen negatif sedangkan nilai tertinggi parameter *recall* diperoleh sentimen netral. Nilai akurasi rata-rata prediksi sebesar 0,72.

Tabel 8. Hasil pengujian model dengan dataset hamburger

No	Nama Label	Parameter		
		Precision	Recall	F1-score
1	Positif	0,67	0,71	0,69
2	Negatif	0,87	0,69	0,77
3	Netral	0,66	0,75	0,7

Kinerja hasil prediksi model dengan dataset spaghetti dideskripsikan dalam Tabel 9 berikut. Nilai *precision* tertinggi dicapai sentimen negatif sedangkan nilai tertinggi parameter *recall* diperoleh sentimen positif. Kemudian, nilai tertinggi *f1-score* diperoleh label sentimen netral. Nilai akurasi rata-rata prediksi sebesar 0,63.

Tabel 9. Hasil pengujian model dengan dataset spaghetti

No	Nama Label	Parameter		
		Precision	Recall	F1-score
1	Positif	0,55	0,75	0,64

No	Nama Label	Parameter		
		Precision	Recall	F1-score
2	Negatif	0,86	0,39	0,54
3	Netral	0,64	0,72	0,68

Secara umum, nilai *precision* tertinggi diperoleh model dataset hamburger pada sentimen negatif. Kemudian, untuk parameter *recall*, nilai tertinggi dicapai model dataset sate pada sentimen positif. Sementara itu, nilai tertinggi pada parameter *f1-score* diperoleh model dataset hamburger pada sentimen negatif.

3.3 Perbandingan Kinerja Model Klasifikasi

Perbandingan kinerja antar model dilakukan dengan menghitung nilai rata-rata parameter *accuracy*, *precision*, *recall* dan *f1-score* setiap model. Perbandingan kinerja setiap model dideskripsikan dalam Tabel 10 dan divisualisasikan dalam grafik bar pada Gambar 6 berikut.

Tabel 10. Perbandingan kinerja model klasifikasi

No	Dataset	Parameter			
		Accuracy	Precision	Recall	F1-score
1	Rendang	0,68	0,69	0,68	0,67
2	Sate	0,68	0,7	0,67	0,67
3	Gudeg	0,68	0,66	0,64	0,64
4	Pizza	0,58	0,64	0,58	0,56
5	Hamburger	0,72	0,72	0,68	0,62
6	Spaghetti	0,63	0,68	0,62	0,62



Gambar 6. Grafik bar perbandingan kinerja model

Berdasarkan Tabel 10 dan Gambar 6 diatas, diperoleh informasi bahwa kinerja model dengan dataset hamburger memiliki kinerja tertinggi pada parameter *accuracy*, *precision* dan *recall* sedangkan model dengan dataset rendang dan sate mencapai nilai tertinggi pada parameter *f1-score*. Sementara itu, model dengan dataset pizza memiliki nilai terendah pada semua parameter.

4. KESIMPULAN

Sentimen masyarakat terhadap makanan tradisional dan internasional dengan data dari twitter telah dianalisis. Hasilnya menunjukkan bahwa makanan gudeg memperoleh persentase sentimen positif tertinggi sebesar 57,9. Sementara itu, makanan rendang mendapat persentase sentimen negatif tertinggi sebesar 21,9. Persentase tertinggi pada sentimen netral diperoleh makanan hamburger sebesar 60,9. Model klasifikasi sentimen berbasis algoritma Naïve bayes telah dibangun. Penelitian ini memiliki 6 model yang mewakili setiap jenis makanan. Selanjutnya, pengujian model dilakukan. Kinerja model dievaluasi dengan pendekatan *confusion matrix*. Perbandingan hasil kinerja tiap model menunjukkan bahwa model dengan dataset hamburger memiliki kinerja terbaik pada variabel *accuracy*, *precision* dan *recall* dengan nilai masing-masing sebesar 0,72, 0,72, dan 0,68. Kemudian, hasil tertinggi pada parameter *f1-score* dicapai oleh model dengan dataset rendang dan sate sebesar 0,67. Sementara itu, model dengan dataset pizza memperoleh nilai terendah pada semua parameter. Berdasarkan hasil evaluasi diatas, kinerja model belum optimal dalam melakukan klasifikasi. Untuk penelitian selanjutnya, *library* sastrawi akan diimplementasikan pada proses *stemming*. Kemudian, kinerja prediksi model klasifikasi perlu ditingkat dengan mengimplementasikan metode *deep learning*. Selain itu, model klasifikasi akan dikembangkan untuk mengidentifikasi semua jenis makanan yang ada.

REFERENCES

- [1] P. Carracedo, R. Puertas, and L. Marti, "Research lines on the impact of the COVID-19 pandemic on business. A text mining analysis," *J. Bus. Res.*, vol. 132, no. September 2020, pp. 586–593, 2021, doi: 10.1016/j.jbusres.2020.11.043.
- [2] S. Mansour, "Social media analysis of user's responses to terrorism using sentiment analysis and text mining," *Procedia Comput. Sci.*, vol. 140, pp. 95–103, 2018, doi: 10.1016/j.procs.2018.10.297.
- [3] Y. Sun *et al.*, "Residents' sentiments towards electricity price policy: Evidence from text mining in social media," *Resour. Conserv. Recycl.*, vol. 160, no. April, p. 104903, 2020, doi: 10.1016/j.resconrec.2020.104903.
- [4] M. Qorib, T. Oladunni, M. Denis, E. Ososanya, and P. Cotae, "Covid-19 Vaccine Hesitancy: Text Mining, Sentiment Analysis and Machine Learning on COVID-19 Vaccination Twitter Dataset," *Expert Syst. Appl.*, vol. 212, no. January 2022, p. 118715, 2022, doi: 10.1016/j.eswa.2022.118715.
- [5] J. Chen, N. Song, Y. Su, S. Zhao, and Y. Zhang, "Learning user sentiment orientation in social networks for sentiment analysis," *Inf. Sci. (Nij.)*, no. xxxx, 2022, doi: 10.1016/j.ins.2022.10.135.
- [6] W. Yulita *et al.*, "Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid-19 Menggunakan Algoritma Naïve Bayes Classifier," *Jdmsi*, vol. 2, no. 2, pp. 1–9, 2021.
- [7] R. Watrighthos, M. Giatman, W. Simatupang, R. Syafriyati, and N. K. Daulay, "Analisis Sentimen Pembelajaran Campuran Menggunakan Twitter Data," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 166, 2022, doi: 10.30865/mib.v6i1.3383.
- [8] M. A. Amrustian, W. Widayat, and A. M. Wirawan, "Analisis Sentimen Evaluasi Terhadap Pengajaran Dosen di Perguruan Tinggi Menggunakan Metode LSTM," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 535, 2022, doi: 10.30865/mib.v6i1.3527.
- [9] B. Laurensz and Eko Sedyono, "Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 2, pp. 118–123, 2021, doi: 10.22146/jnteti.v10i2.1421.
- [10] A. L. Fairuz, R. D. Ramadhani, and N. A. F. Tanjung, "Analisis Sentimen Masyarakat Terhadap COVID-19 Pada Media Sosial Twitter," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 1, no. 1, pp. 42–51, 2021, doi: 10.20895/dinda.v1i1.180.
- [11] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.
- [12] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 10, no. 1, pp. 34–40, 2022, doi: 10.23960/jitet.v10i1.2262.
- [13] T. Rivanie, R. Pebrianto, T. Hidayat, A. Bayhaqy, W. Gata, and H. B. Novitasari, "Analisis Sentimen Terhadap Kinerja Menteri Kesehatan Indonesia Selama Pandemi Covid-19," *J. Inform.*, vol. 21, no. 1, pp. 1–13, 2021, doi: 10.30873/ji.v21i1.2864.
- [14] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification," *Sistemasi*, vol. 10, no. 1, p. 222, 2021, doi: 10.32520/stmsi.v10i1.1179.
- [15] A. Rahman Isnain, A. Indra Sakti, D. Alita, and N. Satya Marga, "Sentimen Analisis Publik Terhadap Kebijakan Lockdown Pemerintah Jakarta Menggunakan Algoritma Svm," *Jdmsi*, vol. 2, no. 1, pp. 31–37, 2021, [Online]. Available: <https://t.co/NhmfMjtXw>
- [16] T. N. Wijaya, R. Indriati, and M. N. Muzaki, "Analisis Sentimen Opini Publik Tentang Undang-Undang Cipta Kerja Pada Twitter," *Jambura J. Electr. Electron. Eng.*, vol. 3, no. 2, pp. 78–83, 2021, doi: 10.37905/jjee.v3i2.10885.
- [17] N. L. P. M. Putu, Ahmad Zuli Amrullah, and Ismarmiaty, "Analisis Sentimen dan Pemodelan Topik Pariwisata Lombok Menggunakan Algoritma Naive Bayes dan Latent Dirichlet Allocation," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, pp. 123–131, 2021, doi: 10.29207/resti.v5i1.2587.
- [18] S. M. Salsabila, A. Alim Murtopo, and N. Fadhilah, "Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naive Bayes Classifier," *J. Minfo Polgan*, vol. 11, no. 2, pp. 30–35, 2022, doi: 10.33395/jmp.v11i2.11640.
- [19] L. Li *et al.*, "Naive Bayes classifier based on memristor nonlinear conductance," *Microelectronics J.*, vol. 129, no. September, 2022, doi: 10.1016/j.mejo.2022.105574.
- [20] H. Mustofa and A. A. Mahfudh, "Klasifikasi Berita Hoax Dengan Menggunakan Metode Naive Bayes," *Walisongo J. Inf. Technol.*, vol. 1, no. 1, p. 1, 2019, doi: 10.21580/wjit.2019.1.1.3915.
- [21] A. D. Wibisono, S. Dadi Rizkiono, and A. Wantoro, "Filtering Spam Email Menggunakan Metode Naive Bayes," *TELEFORTECH J. Telemat. Inf. Technol.*, vol. 1, no. 1, pp. 9–17, 2020, doi: 10.33365/tft.v1i1.685.
- [22] F. Xu, Z. Pan, and R. Xia, "E-commerce product review sentiment classification based on a naïve Bayes continuous learning framework," *Inf. Process. Manag.*, vol. 57, no. 5, p. 102221, 2020, doi: 10.1016/j.ipm.2020.102221.
- [23] S. Khairunnisa, A. Adiwijaya, and S. Al Faraby, "Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19)," *J. Media Inform. Budidarma*, vol. 5, no. 2, p. 406, 2021, doi: 10.30865/mib.v5i2.2835.
- [24] R. N. Fahmi, Nursyifa, and A. Primajaya, "Analisis Sentimen Pengguna Twitter Terhadap Kasus Penembakan Laskar Fpi Oleh Polri Dengan Metode Naive Bayes Classifier," *JIKO (Jurnal Inform. dan Komputer)*, vol. 5, no. 2, pp. 61–66, 2021, [Online]. Available: <https://ejournal.akakom.ac.id/index.php/jiko/article/view/437>
- [25] R. Asmara, "Analisis Sentimen Temporal tentang Kuliner di Kota Surabaya Berbasis Gender menggunakan Bahasa Indonesia," vol. 5, no. 1, 2020.
- [26] R. N. Fahmi, T. Informatika, U. Singaperbangsa, and T. Timur, "Resti Noor Fahmi 1, Nursyifa 2, dan Aji Primajaya 3 Teknik Informatika Universitas Singaperbangsa Karawang Jl. HS.Ronggo Waluyo, Telukjambe Timur, Karawang," vol. 5, no. 2, pp. 61–66, 2021.
- [27] N. B. Classifier and N. Bayes, "ANALISIS SENTIMEN HASHTAG KULINER DI INDONESIA," vol. 2, no. 1, pp. 19–25, 2021.



- [28] I. A. Mastan and Y. Toni, “MENGUNAKAN METODE NAÏVE BAYES CLASSIFIER SENTIMENT ANALYSIS OF CHICKEN CULINARY PLACE FROM VISITORS COMMENTS USING NAÏVE BAYES CLASSIFIER METHOD,” vol. 3, no. 1, pp. 42–50, 2020.
- [29] A. Dan and N. Bayes, “Analisis sentimen terhadap,” vol. 3, pp. 163–174, 2017.
- [30] I. T. Julianto, “Analisis Sentimen Terhadap Sistem Informasi Akademik Institut Teknologi Garut,” *J. Algoritma*, vol. 19, no. 1, pp. 449–456, 2022, doi: 10.33364/algoritma/v.19-1.1112.
- [31] K. Jindal and R. Aron, “A systematic study of sentiment analysis for social media data,” *Mater. Today Proc.*, no. xxxx, 2021, doi: 10.1016/j.matpr.2021.01.048.
- [32] N. Azhar, P. P. Adikara, S. Adinugroho, and P. Korespondensi, “Analisis Sentimen Ulasan Kedai Kopi Menggunakan Metode Naïve Bayes Dengan Seleksi Fitur Algoritme Genetika Sentiment Analysis for Coffee Shop Reviews Using Naïve Bayes Method With Genetic Algorithm Feature Selection,” vol. 8, no. 3, pp. 609–618, 2021, doi: 10.25126/jtiik.202184436.

Paper JSON

ORIGINALITY REPORT

20%

SIMILARITY INDEX

19%

INTERNET SOURCES

5%

PUBLICATIONS

3%

STUDENT PAPERS

PRIMARY SOURCES

1	ejournal.stmik-budidarma.ac.id Internet Source	11%
2	Submitted to Universitas Pamulang Student Paper	3%
3	inacl.id Internet Source	1%
4	www.researchgate.net Internet Source	1%
5	id.123dok.com Internet Source	1%
6	Submitted to Universitas Negeri Jakarta Student Paper	<1%
7	jurnal.unej.ac.id Internet Source	<1%
8	adoc.tips Internet Source	<1%
9	farcoraharjo.blogspot.com Internet Source	<1%

10	jtiik.ub.ac.id Internet Source	<1 %
11	nemertes.library.upatras.gr Internet Source	<1 %
12	Hasim Budi Jatmiko, Nanang Tedi Kurniadi, Dony Maulana. "Optimasi Naïve Bayes Dengan Particle Swarm Optimization Untuk Analisis Sentimen Formula E-Jakarta", Journal Automation Computer Information System, 2022 Publication	<1 %
13	Kamaram H. Manguri, Rebaz N. Ramadhan, Pshko R. Mohammed Amin. "Twitter Sentiment Analysis on Worldwide COVID-19 Outbreaks", Kurdistan Journal of Applied Research, 2020 Publication	<1 %
14	ejnteti.jteti.ugm.ac.id Internet Source	<1 %
15	ejournal.akprind.ac.id Internet Source	<1 %
16	herinazulkarneni.wordpress.com Internet Source	<1 %
17	repository.uin-suska.ac.id Internet Source	<1 %

18	Internet Source	<1 %
19	www.coursehero.com Internet Source	<1 %
20	www.ilmiahpendidikan.com Internet Source	<1 %
21	www.journal.uad.ac.id Internet Source	<1 %
22	Miftahul Qorib, Timothy Oladunni, Max Denis, Esther Ososanya, Paul Cotae. "Covid-19 Vaccine Hesitancy: Text Mining, Sentiment Analysis and Machine Learning on COVID-19 Vaccination Twitter Dataset", Expert Systems with Applications, 2022 Publication	<1 %

Exclude quotes On

Exclude matches Off

Exclude bibliography On