

Analisis Sentimen terhadap Opini Feminisme Menggunakan Metode Naive Bayes

Widya Wahyuni^{1✉}

¹Independent Researcher

widyawahyuni1997@gmail.com

Abstract

The perspective of the development of feminism centered on women around the world who wants to be free from pressure, oppression and inequality from men, continues to this day. Various public opinions about feminism are now contained in various social media. Long debates about criticism and support for feminism in equalizing women's position both in terms of intellect, and the role of women in making decisions. This research was conducted with the aim of looking at public sentiment based on opinions circulating on social media. Hashtags or hash tags related to feminism from social media are the main data that will be used to analyze public opinion sentiment about feminism and 600 data are obtained about feminism. The data obtained were separated into positive, negative and neutral opinions for analysis using Naïve Bayes (NB). The results of using the Naïve Bayes method obtained a recall value of 84%, precision 94% and Fi-Score of 86% with an accuracy of 88%. Through this research, the results of classification using the Naive Bayes method in analyzing sentiment against feminist opinions have good performance.

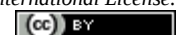
Keywords: Feminism, Sentiment Analysis, Support Vector Machine, Naive Bayes, Opinion.

Abstrak

Perspektif perkembangan feminisme yang berpusat pada perempuan diseluruh dunia yang berkeinginan untuk lepas dari tekanan, penindasan dan ketimpangan dari kaum laki-laki, terus berlanjut hingga sekarang. Berbagai opini publik mengenai feminisme kini tertuang dalam berbagai media sosial. Perdebatan panjang mengenai kritikan hingga dukungan pada feminisme dalam menyamakan kedudukan perempuan baik dari segi intelektualitas, dan peran perempuan dalam mengambil keputusan. Penelitian ini dilakukan dengan tujuan untuk melihat sentimen publik berdasarkan opini-opini yang beredar di media sosial. Tagar atau hash tag yang berkaitan dengan feminisme dari media sosial menjadi data utama yang akan digunakan untuk menganalisis sentimen opini publik mengenai feminisme dan didapatkan 600 data mengenai feminisme. Data yang diperoleh tersebut dipisahkan menjadi opini positif, negatif dan netral untuk dianalisa menggunakan Naïve Bayes (NB). Hasil dari penggunaan metode Naïve Bayes diperoleh nilai recall 86%, precision 89% dan Fi-Score sebesar 86% dengan hasil accuracy sebesar 87%. Melalui penelitian ini didapatkan hasil klasifikasi menggunakan metode Naive Bayes dalam melakukan analisa sentimen terhadap opini feminisme memiliki performa yang baik.

Kata kunci: Feminisme, Analisis Sentimen, Support Vector Machine, Naïve Bayes, Opini.

INFEB is licensed under a Creative Commons 4.0 International License.



1. Pendahuluan

Mayoritas penggunaan media sosial digunakan untuk mengungkapkan pendapat atau pandangan tentang suatu topik, produk atau berita [1]. Setiap hari pengguna media sosial terus meningkat [2]. Tercatat pada tahun 2019, penggunaan media sosial diperkirakan naik hingga 2,77 miliar pengguna di dunia. Beragam tipe informasi di upload dan dibagikan di media sosial baik dalam bentuk teks, video, foto dan audio [3]. Kebanyakan penggunaan media sosial digunakan untuk mengekspresikan perasaan, keseharian hingga pendapat mengenai berbagai macam hal [4]

Besarnya perkembangan Web 2.0 telah merubah cara seorang individu dalam mengekspresikan perspektif dan perasaan seseorang. Mulai dari komentar, pendapat, usul, pandangan bahkan hingga kritik yang

berkaitan dengan suatu topik tertentu dibagikan dalam jejaring sosial online [5]. Salah satu media sosial yang banyak dipakai adalah twitter [6]. Twitter banyak digunakan untuk menyampaikan keluhan, kesah dan opini masyarakat terhadap suatu pemikiran ataupun layanan [7]. Salah satunya adalah feminisme. Feminisme percaya dalam hal kebijakan publik dan sikap tentang gender [8].

Feminisme pertama kali muncul di Prancis karena tradisi turun menurun yang seringkali merendahkan kaum Wanita [9]. Feminisme tak lepas dari kontra dan kritik mengenai perkembangannya, seperti perbedaan pendapat tokoh feminis dengan kajian Islam. Sebuah pemikiran dari Amina Wadud tokoh feminis yang menafsirkan ayat al-Qur'an melalui kacamata feminis beranggapan bahwa al-Qur'an yang ditafsirkan oleh ulama klasik mengusung bias patriarki [10]. Feminisme telah menghasilkan banyak pencapaian

emansipatoris yang luar biasa. Berhasil mengubah hukum dan praktik kelembagaan dalam standar diskriminasi gender, pekerjaan dan partisipasi perempuan di ruang public [11].

Menurut pandangan feminis, gender tidak seharusnya dipahami sebagai konstruksi sosial, tetapi juga sebagai relasional dan performatif, sehingga mungkin untuk memahami kerentanan tidak hanya terjadi pada perempuan [12]. Teori feminis mengungkapkan tentang kekuasaan beroperasi berdasarkan sistem sosial dan hierarki [13].

Data memungkinkan untuk menafsirkan aspek seseorang pada topik tertentu dan informasi yang dapat dimanfaatkan untuk melakukan prediksi pada suatu topik yang berguna untuk mengambil keputusan yang tepat berdasarkan pertimbangan analisis sentimen [14]. Analisis sentimen dapat digunakan untuk mengekstrak sentimen dan pendapat dari teks [15].

Analisis Sentimen dapat mendeteksi polaritas suatu opini kedalam opini positif atau negatif dalam suatu teks, yang dapat berupa dokumen, paragraf, kalimat ataupun klausa [16]. Analisis sentimen tidak hanya digunakan untuk kepentingan bisnis, informasi yang tersedia dapat digunakan untuk kepentingan sosial, Pendidikan ataupun hiburan [17]. Sistem dilatih berdasarkan kumpulan data pelatihan, dan setelah pelatihan, sistem mencoba membuat klasifikasi otomatis atau dapat membuat prediksi. Efisiensi sistem diuji berdasarkan nilai prediksi yang benar dan kumpulan data uji [18].

Proses dari analisis sentimen pada penelitian ini menggunakan metode *Naïve Bayes*. Metode *Naïve Bayes* merupakan algoritma pemantauan *machine learning* yang mendasar dan terkenal berdasarkan teorema *Bayesian* dan anggapan kondisi karakteristik otonom. Metode ini mendalilkan aksesibilitas pada tahap dokumen dengan tampilan dan label yang telah ditentukan sebelumnya [19]. *Naïve Bayes* berasal dari dua suku kata, yaitu *Naïve* yang berasal dari asumsi kemunculan suatu fitur adalah independent dari terjadinya fitur lain sehingga setiap fitur berkontribusi secara individual untuk klasifikasi tanpa ketergantungan dengan fitur lain. Sedangkan *bayes* berasal dari prinsip teorema bayes [20]. Keuntungan menggunakan *naïve bayes* untuk mengklasifikasikan salah satunya karena untuk mengklasifikasikan menggunakan jumlah data pelatihan yang kecil [21].

Metode SVM dan *Naïve Bayes* digunakan dalam penelitian untuk melihat sentimen terhadap MyIndiHome. Penelitian ini menggunakan data yang dikumpulkan dalam rentang waktu 1 November 2020 hingga 15 Desember 2020 dengan total 2.539, diperoleh dari aplikasi *Google Play*. Hasil dari penelitian ini menunjukkan bahwa tingkat *error* yang terjadi pada Indihome masih terbilang tinggi. Rata-rata total akurasi metode SVM sebesar 86,54% lebih baik dari *Naïve bayes* [22].

Analisis Sentimen terhadap Opini mengenai vaksinasi Covid-19 menggunakan metode *Naïve Bayes* dan SVM. Penelitian dilakukan untuk melihat opini masyarakat mengenai Covid-19, di mana pada saat itu masyarakat banyak menuangkan opininya ke sosial media. Penelitian ini menggunakan 1000 data yang berkisar dari tahun 2020 yang berasal dari twitter. Hasil dari penelitian ini untuk melihat perbandingan antara *Naïve Bayes* dan SVM. SVM memiliki performa yang lebih baik dari pada *Naïve Bayes* [23].

Analisis sentimen juga dapat digunakan untuk menganalisis sentimen terhadap politik, seperti penelitian yang dilakukan Akhmad Mazuki dan Arita Witanti yang meneliti tentang analisis sentiment masyarakat di *twitter*. Penelitian ini menggunakan 1200 yang diambil dari *twitter* yang kemudian akan dianalisis menggunakan metode *Naïve Bayes* dan *Support Vector Machine*. Hasil analisis yang dilakukan didapatkan hasil bahwa *penggunaan Naïve Bayes* ternyata lebih baik ketimbang *Support Vector Machine* dengan tingkat performansi untuk klasifikasi *tweet* 92,2% [24].

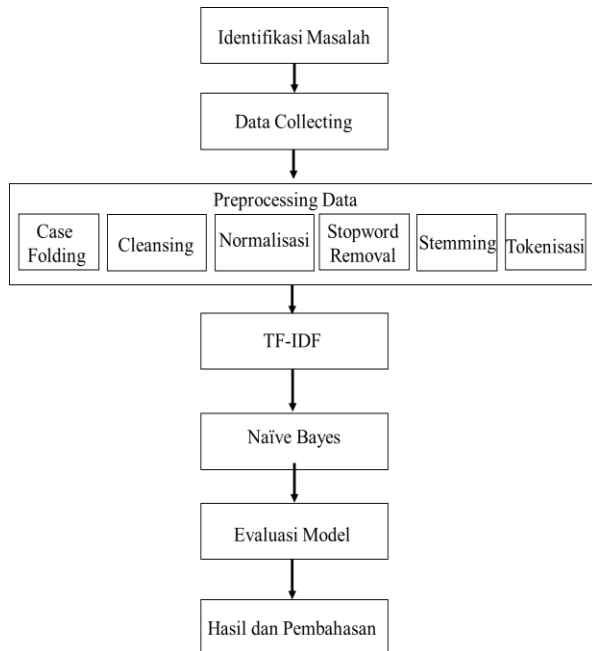
Melihat tingkat kepuasan konsumen terhadap suatu aplikasi dapat menggunakan analisis sentiment, seperti penelitian yang membahas mengenai sentimen pengguna terhadap aplikasi Ruangguru. Analisis ini menggunakan dataset sebanyak 1629 dengan bentuk text review dalam rentang waktu 01 Maret sampai 30 Maret 2020. Menggunakan metode *Naïve Bayes*, *Random Forest* dan *Support Vector Machine* dataset akan diuji. Berdasarkan hasil Analisa dan hasil pengujian dari beberapa metode didapatkan hasil akurasi tertinggi pada metode *Random Forest* dengan jumlah akurasi 97,16%, *Support Vector Machine* dengan akurasi 96,01%, dan nilai akurasi terendah pada pengujian *Naïve Bayes* dengan nilai akurasi 94,16%. Kesimpulan dari penelitian ini adalah berdasarkan tabel hasil *preprocessing* dapat disimpulkan bahwa komentar positif lebih banyak ketimbang komentar negatif [25].

Penelitian ini dilakukan karena perkembangan dari feminisme yang merupakan paham yang datang dari barat, telah berkembang di Indonesia. Perkembangan ini juga yang mengubah pola pikir perempuan, Feminisme menimbulkan pro dan kontra dikalangan masyarakat. Karena pro dan kontra maka untuk melihat tingkat sentiment opini masyarakat terhadap feminisme dilakukan analisis sentimen, di mana sentimen dari masyarakat dapat dipisahkan menjadi sentimen positif, netral dan negatif.

Hasil dari penelitian ini nantinya menunjukkan tingkat akurasi dari analisis sentimen yang dilakukan menggunakan algoritma *Support Vector Machine* dan *Naïve Bayes*.

2. Metodologi Penelitian

Penelitian ini dilakukan secara bertahap dan terstruktur di mana pada setiap tahapannya dilakukan menggunakan metode yang kritis dan mampu merahkan pada proses yang tepat. Metode penelitian ini adalah Teknik-teknik yang digunakan dalam merumuskan, menganalisa, mengumpulkan data sampai proses implementasi hasil penelitian. Metode penelitian disajikan pada Gambar 1.



Gambar 1. Metode Penelitian

Penelitian ini menggunakan dataset yang berasal dari twitter dengan *query* yang berkaitan dengan feminisme. Hasil data yang diperoleh yaitu sebanyak 600 data yang dibagi menjadi 80% data latih sejumlah 400 data dan 20% data uji sejumlah 200 data.

3. Hasil dan Pembahasan

1. Identifikasi Masalah

Kualitas penelitian dipengaruhi oleh masalah penelitian, sebagai upaya untuk menjelaskan masalah dan dapat mengukur penjelasan masalah tersebut identifikasi masalah dilakukan. Jadi, identifikasi masalah merupakan pendefinisian dari masalah yang ada dalam sebuah penelitian.

2. Data Collecting

Pengumpulan data dilakukan dengan Teknik *crawling* data. Data yang diambil berasal dari *platform Twitter* dengan *query* yang berkaitan dengan feminisme. Setelah dilakukan *crawling* data, diperoleh 600 record data dari periode 27 Juli hingga 8 Agustus 2022. Data yang diperoleh masih dalam bentuk yang tidak terstruktur, sehingga tidak bisa langsung digunakan.

3. Case Folding

Case folding bertujuan mengubah semua bentuk huruf dalam sebuah teks menjadi huruf kecil untuk mendapatkan data yang akurat pada proses selanjutnya. Contoh hasil dari proses *case folding* disajikan pada Tabel 1.

Tabel 1. Contoh Hasil *Case Folding*

Data Input	Data Output
RT @Dipoett: @neveurlandx @infotwitwor_ Sama aja kok kak kadang cewek feminis garis keras ya goblok, udah mending moderate aja	rt @dipoett: @neveurlandx @infotwitwor_ sama aja kok kak kadang cewek feminis garis keras ya goblok, udah mending moderate aja

4. Cleansing

Pada tahapan ini dilakukan suatu proses untuk mendeteksi dan memperbaiki record yang *corrupt* atau tidak akurat. Tahapan ini penting karena data yang kotor akan mempengaruhi tingkat akurasi dari analisis maupun prediksi yang akan dilakukan. Sebelum memasuki tahap untuk membersihkan data yang dinilai missing value, data akan dilihat terlebih dahulu mengenai deskripsi atau info terkait dataset lebih lanjut. Hasil dari proses cleansing disajikan pada Gambar 2.

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 599 entries, 0 to 598
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype
---  ---
 0   Tanggal    599 non-null    object
 1   User       599 non-null    object
 2   Tweet      599 non-null    object
 3   Id         599 non-null    float64
 4   Label      599 non-null    object
dtypes: float64(1), object(4)
memory usage: 23.5+ KB
  
```

Gambar 2. Hasil Proses Cleansing

5. Normalisasi

Normalisasi merupakan tahap untuk menghilangkan sejumlah tanda baca dan, angka, simbol, URL dan username agar data yang didapat dapat lebih terstruktur untuk proses selanjutnya. Contoh hasil dari proses Normalisasi disajikan pada Tabel 3.

Tabel 3. Contoh Hasil Normalisasi

Data Input	Data Output
rt @dipoett: @neveurlandx @infotwitwor_ sama aja kok kak kadang cewek feminis garis keras ya goblok, udah mending moderate aja	sama aja kok kak kadang cewek feminis garis keras ya goblok, udah mending moderate aja

6. Stopword Removal

Pada tahapan ini dilakukan menghilangkan frasa yang tidak penting sesuai dengan frasa yang terdapat pada stopwords. Jika terdapat kata yang sesuai dengan kata yang terdapat pada stopwords maka kata tersebut akan dihapus dari dokumen karena dianggap sebagai kata

yang tidak penting. Contoh hasil *stopword removal* disajikan pada Tabel 4.

Tabel 4. Contoh Hasil *Stopword Removal*

Data Input	Data Output
sama aja kok kak kadang cewek feminis garis keras ya goblok, udah mending moderate aja	aja kak kadang cewek feminis garis keras ya goblok, udah mending moderate aja

7. Stemming

Proses stemming dilakukan untuk menghilangkan setiap kata imbuhan pada kata baik yang ada pada awal kata atau akhir kata dengan memperhatikan awalan umum dan sufiks, yang dapat ditemukan dalam kata. Contoh dari hasil proses *stemming* disajikan pada Tabel 5.

Tabel 5. Contoh Hasil Stemming

Data Input	Data Output
aja kak kadang cewek feminis garis keras ya goblok, udah mending moderate aja	aja kak kadang cewek feminis gari kera ya goblok udah mending moderat aja

8. Tokenisasi

Pada tahapan ini akan menghilangkan tanda baca yang tidak diperlukan dan memotong teks menjadi kata, simbol, karakter atau tanda baca, sehingga menjadi token yang dapat dianalisis. Contoh hasil proses tokenisasi disajikan pada Tabel 6.

Tabel 6. Contoh Hasil Tokenisasi

Data Input	Data Output
aja kak kadang cewek feminis garis keras ya goblok, udah mending moderat aja	'aja', 'kak', 'kadang', 'cewek', 'feminis', 'gari', 'kera', 'ya', 'goblok', 'udah', 'mending', 'moderat', 'aja'

2.1 Ekstraksi Fitur

Pada proses ekstraksi fitur akan ditentukan word vector dari dataset yang telah dilakukan proses tokenisasi. Sebagai contoh dari 600 record yang telah dikumpulkan disajikan 2 data dari 600 record. Contoh dari record yang digunakan untuk sampel disajikan pada Tabel 7.

Tabel 7. Data Sampel

DOC	Data Output
D1	'aja', 'kak', 'kadang', 'cewek', 'feminis', 'gari', 'kera', 'ya', 'goblok', 'udah', 'mending', 'moderat', 'aja'
D2	'topik', 'utama', 'tgg', 'sehat', 'mental', 'topik', 'yg', 'angkat', 'feminisme', 'percaya', 'stigma', 'masyarakat', 'parenting', 'buku', 'jg', 'nyaman', 'bgt', 'baca', 'pilih', 'bahasa', 'ringan', 'kecuali', 'sain', 'kawan', 'alir', 'gitu', 'aja'
D3	'bilang', 'gitu', 'beliau', 'tetep', 'dukung', 'nove', 'perempuan', 'feminis', 'abi', 'bilang', 'cerita', 'keren', 'buku', 'apa', 'breast', 'and', 'egg', 'karya', 'kawakami', 'mieko'

Tabel 12. Tabel Hasil TF-IDF

Term	TF			DF	D/DF	IDF	IDF+1	TF-IDF		
	D1	D2	D3					D1	D2	D3
Feminis	1	0	1	2	1.5	0,176	1,176	1,176	0	1,176
Goblok	1	0	0	1	3	0,477	1,477	1,477	0	0
Feminisme	0	1	0	1	3	0,477	1,477	0	1,477	0
Masyarakat	0	1	0	1	3	0,477	1,477	0	1,477	0
Perempuan	0	0	1	1	1.5	0,176	1,176	0	0	1,176
Nilai Bobot Setiap Dokumen								2.653	2.954	2.352

Berdasarkan Tabel 7 terdapat beberapa kata yang dipilih sebagai kata kunci yang akan digunakan, yaitu feminis, goblok, feminisme, masyarakat, dan perempuan.

Pada Tabel 7 disajikan sebuah *vector* yang merupakan perwujudan dari setiap dokumen, di mana kata setiap kata dalam dokumen diberi nilai satu jika kata tersebut terdapat dalam dokumen tersebut, dan 0 untuk kata yang tidak terdapat pada dokumen yang disajikan pada Tabel 8.

Tabel 8. *Word Vector*

DOC	Feminis	Goblok	Feminisme	masyarakat	Perempuan
D1	1	1	0	0	0
D2	0	0	1	1	0
D3	1	0	0	0	1

Hasil dari word vector kemudian dihitung menggunakan TF-IDF sehingga akan menghasilkan nilai *word vector* yang sudah terbobot. Proses pertama yaitu menghitung jumlah term yang muncul dalam semua dokumen. Hasil dari nilai DF disajikan pada Tabel 9.

Tabel 9. Tabel DF

Term	DF (Document Frequency)
Feminis	2
Goblok	1
Feminisme	1
Masyarakat	1
Perempuan	1

Hasil dari DF digunakan untuk menghitung IDF (*Inverse Document Frequency*) di mana nilai dari IDF disajikan dalam Tabel 10. Rumus untuk menentukan nilai IDF disajikan dalam Persamaan (1).

$$idf_t = \log_{10} \frac{N}{df_t} \quad (1)$$

Tabel 10. Tabel Nilai IDF

Term	DF	D/DF	IDF
Feminis	2	1.5	0,176
Goblok	1	3	0,477
Feminisme	1	3	0,477
Masyarakat	1	3	0,477
Perempuan	1	3	0,477

Setelah nilai IDF didapatkan selanjutnya mencari nilai TF-IDF. Hasil dari nilai TF-IDF disajikan dalam Tabel 12. Rumus untuk menghasilkan TF-IDF disajikan dalam Persamaan (2) yang disajikan Tabel 11.

$$w_{t,d} = w_{t,f,t,d} \times (idf_t + 1) \quad (2)$$

Nilai yang disajikan pada Tabel 12 kemudian akan disajikan ke dalam bentuk *word vector* yang sudah terbobot. Nilai dari *word vector* yang sudah terbobot disajikan dalam Tabel 13.

Tabel 13. Tabel *Word Vector* Terbobot

DOC	Feminis	Goblok	Feminisme	Masyarakat	Perempuan
D1	1,176	1,477	0	0	0
D2	0	0	1,477	1,477	0
D3	1,176	0	0	0	1,176

2.2 Klasifikasi *Naïve Bayes*

Naïve Bayes menggunakan model probabilistic untuk proses klasifikasi berdasarkan teorema bayes. Klasifikasi *naïve bayes* menggunakan rumus yang disajikan dalam Persamaan (3), Persamaan (4), dan Persamaan (5).

$$P(c) = \frac{N_c}{N} \quad (3)$$

$$P(w|c) = \frac{\text{count}(c,w) + a_w}{\text{count}(c) + a} \quad (4)$$

$$P(c|w_i) = \argmin P(c) \prod_{i=1}^n \frac{1}{P(w_i|c)^{f_i}} \quad (5)$$

Dimana Persamaan (3), Persamaan (4) dan Persamaan (5), w merupakan kata, dan c merupakan kelas target. N_c merupakan jumlah kelas c . N adalah jumlah seluruh kelas. $\text{Count}(c,w)$ menunjukkan berapa kali w muncul selain kelas c . $\text{Count}(c)$ adalah jumlah dari seluruh kata yang muncul dari selain kelas c . a_w merupakan smoothing parameter diberi nilai 1. a adalah jumlah keseluruhan dari a_w .

Data yang digunakan untuk menentukan kelas sentimen adalah data uji. Hal ini dilakukan setelah dihasilkan nilai prior dari data latih.

$$P(X|\text{Positif}) = P(\text{Positif}) \times P(w|\text{Positif})$$

$$P(X|\text{Positif}) = 0,6 \times 0,00000003061$$

$$P(X|\text{Positif}) = 0,00000001837$$

$$P(X|\text{Negatif}) = P(\text{Negatif}) \times P(w|\text{Negatif})$$

$$P(X|\text{Negatif}) = 0,4 \times 0,00000007616$$

$$P(X|\text{Negatif}) = 0,0000000457$$

$$P(X|\text{Netral}) = P(\text{Netral}) \times P(w|\text{Netral})$$

$$P(X|\text{Netral}) = 0,2 \times 0,00000008804$$

$$P(X|\text{Netral}) = 0,00000005282$$

Berdasarkan perhitungan dari tiap term pada data uji nilai tertinggi terdapat pada kelas netral. Sehingga dapat dikatakan bahwa data uji memiliki klasifikasi kelas netral.

2.3 Evaluasi Model

Sistem yang dihasilkan dilakukan untuk mengetahui seberapa baik performa dari metode yang digunakan. Metode evaluasi yang digunakan adalah accuracy,

precision, recall, dan FI-Score. Nilai evaluasi yang dibandingkan hanya FI-Score saja, karena merupakan metrik yang lebih baik untuk data yang tidak seimbang. Menghitung metode evaluasi membutuhkan True Positive (TP), True Negative (TN), False Positive (FP) dan False Negative (FN) dari confusion matrix setiap label. Rumus dari metode evaluasi disajikan dalam persamaan (6), (7), (8), (9)

$$\text{accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (6)$$

$$\text{precision} = \frac{TP}{TP+FP} \quad (7)$$

$$\text{recall} = \frac{TP}{TP+FN} \quad (8)$$

$$F1 \text{ Score} = \frac{2 \times (\text{recall} \times \text{precision})}{(\text{recall} + \text{precision})} \quad (9)$$

Hasil dari performansi metode *Naïve Bayes* dalam melakukan analisis sentimen mengenai opini masyarakat tentang feminisme disajikan pada Tabel 14.

Tabel 14. Hasil Evaluasi Performansi

Aktual/Prediksi	Precisiom	Recall	F1-Score
Positif Label	80%	97%	88%
Negatif Label	94%	70%	80%
Netral Label	92%	91%	91%
Rata-rata	89%	86%	86%
Accuracy	87%		

4. Kesimpulan

Proses analisis sentimen dilakukan untuk memprediksi suatu fenomena ataupun keperluan bisnis. Pada penelitian ini hasil dari analisis sentimen terhadap opini masyarakat mengenai feminisme akan memberikan gambaran pada masyarakat seberapa tinggi atau rendahnya feminisme berkembang di Indonesia. Selain itu masyarakat bisa dapat memberikan pengetahuan baru mengenai penggunaan analisis sentiment dalam menganalisis tingkat fenomena sosial tertentu yang terjadi pada masyarakat. Hasil performansi evaluasi analisis sentiment terhadap opini masyarakat mengenai feminisme menggunakan metode *Naïve Bayes* memiliki akurasi sebesar 87%, dimana dari data sejumlah 600 data dengan rasio 80:20 terdapat nilai recall 86%, precision 89% dan Fi-Score sebesar 86%. Hasil accuracy sebesar 87% dapat dikatakan baik, sehingga metode *Naïve Bayes* dapat dikatakan memiliki performa yang baik dalam melakukan analisis sentimen opini masyarakat mengenai feminisme.

Daftar Rujukan

- [1] Chauhan, P., Sharma, N., & Sikka, G. (2021). The emergence of social media data and sentiment analysis in election prediction. *Journal of Ambient Intelligence and Humanized Computing*, 12(2), 2601-2627. DOI: <https://doi.org/10.1007/s12652-020-02423-y>
- [2] Hruska, J., & Maresova, P. (2020). Use of social media platforms among adults in the United States—behavior on social media. *Societies*, 10(1), 27. DOI: <https://doi.org/10.3390/soc10010027>

- [3] Drus, Z., & Khalid, H. (2019). Sentiment analysis in social media and its application: Systematic literature review. *Procedia Computer Science*, 161, 707-714. DOI: <https://doi.org/10.1016/j.procs.2019.11.174>
- [4] Sharma, D., Sabharwal, M., Goyal, V., & Vij, M. (2020). Sentiment analysis techniques for social media data: A review. In *First international conference on sustainable technologies for computational intelligence* (pp. 75-90). Springer, Singapore. DOI: https://doi.org/10.1007/978-981-15-0029-9_7
- [5] Alarifi, A., Tolba, A., Al-Makhadmeh, Z., & Said, W. (2020). A big data approach to sentiment analysis using greedy feature selection with cat swarm optimization-based long short-term memory neural networks. *The Journal of Supercomputing*, 76(6), 4414-4429. DOI: <https://doi.org/10.1007/s11227-018-2398-2>
- [6] Sharma, A., & Ghose, U. (2020). Sentimental analysis of twitter data with respect to general elections in india. *Procedia Computer Science*, 173, 325-334. DOI: <https://doi.org/10.1016/j.procs.2020.06.038>
- [7] Zahoor, S., & Rohilla, R. (2020, August). Twitter sentiment analysis using machine learning algorithms: a case study. In *2020 International Conference on Advances in Computing, Communication & Materials (ICACCM)* (pp. 194-199). IEEE. DOI: <https://doi.org/10.1109/ICACCM50413.2020.9213011>
- [8] Elder, L., Greene, S., & Lizotte, M. K. (2021). Feminist and anti-feminist identification in the 21st century United States. *Journal of Women, Politics & Policy*, 42(3), 243-259. DOI: <https://doi.org/10.1080/1554477X.2021.1929607>
- [9] Tamboukou, M. (2020). Women Workers' Education. *Handbook of Historical Studies in Education: Debates, Tensions, and Directions*, 813-829. DOI: https://doi.org/10.1007/978-981-10-2362-0_48
- [10] Rif'at, D. F., & Nurwahidin, N. (2022). Feminisme Dan Kesetaraan Gender Dalam Kajian Islam Kontemporer. *Syntax Literate; Jurnal Ilmiah Indonesia*, 7(1), 172-182. DOI: <http://dx.doi.org/10.36418/syntax-literate.v7i1.6038>
- [11] Alexeyeff, K. (2020, May). Cinderella of the south seas? Virtuous victims, empowerment and other fables of development feminism. In *Women's Studies International Forum* (Vol. 80, p. 102368). Pergamon. DOI: <https://doi.org/10.1016/j.wsif.2020.102368>
- [12] Yadav, P., Saville, N., Arjyal, A., Baral, S., Kostkova, P., & Fordham, M. (2021). A feminist vision for transformative change to disaster risk reduction policies and practices. *International Journal of Disaster Risk Reduction*, 54, 102026. DOI: <https://doi.org/10.1016/j.ijdrr.2020.102026>
- [13] Cannon, C. (2020). On the importance of feminist theories: Gender, race, sexuality and IPV. In *Intimate partner violence and the LGBT+ community* (pp. 37-52). Springer, Cham. DOI: <https://doi.org/10.1016/j.erss.2021.102005>
- [14] Pathak, A. R., Pandey, M., & Rautaray, S. (2021). Topic-level sentiment analysis of social media data using deep learning. *Applied Soft Computing*, 108, 107440. DOI: <https://doi.org/10.1016/j.asoc.2021.107440>
- [15] Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226, 107134. DOI: <https://doi.org/10.1016/j.knosys.2021.107134>
- [16] Priyadarshini, I., & Cotton, C. (2021). A novel LSTM-CNN-grid search-based deep neural network for sentiment analysis. *The Journal of Supercomputing*, 77(12), 13911-13932. DOI: <https://doi.org/10.1007/s11227-021-03838-w>
- [17] Kumar, A., & Garg, G. (2020). Systematic literature review on context-based sentiment analysis in social multimedia. *Multimedia tools and Applications*, 79(21), 15349-15380. DOI: <https://doi.org/10.1007/s11042-019-7346-5>
- [18] Singh, R., & Singh, R. (2021). Applications of sentiment analysis and machine learning techniques in disease outbreak prediction-A review. *Materials Today: Proceedings*. DOI: <https://doi.org/10.1016/j.matpr.2021.04.356>
- [19] Jindal, K., & Aron, R. (2021). A systematic study of sentiment analysis for social media data. *Materials today: proceedings*. DOI: <https://doi.org/10.1016/j.matpr.2021.01.048>
- [20] Soumya, S., & Pramod, K. V. (2020). Sentiment analysis of Villavicencio, C., Macrohon, J. J., Inbaraj, X. A., Jeng, J. H., & Hsieh, J. G. (2021). Twitter sentiment analysis towards covid-19 vaccines in the Philippines using naïve bayes. *Information*, 12(5), 204. DOI: <https://doi.org/10.3390/info12050204>
- [21] Geetha, M. P., & Renuka, D. K. (2021). Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model. *International Journal of Intelligent Networks*, 2, 64-69. DOI: <https://doi.org/10.1016/j.ijin.2021.06.005>
- [22] Hakim, S. N., Putra, A. J., & Khasanah, A. U. (2021). Sentiment analysis on myindihome user reviews using support vector machine and naïve bayes classifier method. *International Journal of Industrial Optimization*, 2(2), 151. DOI: <https://doi.org/10.12928/ijio.v2i2.4437>
- [23] Fitriana, F., Utami, E., & Al Fatta, H. (2021). Analisis Sentimen Opini Terhadap Vaksin Covid-19 pada Media Sosial Twitter Menggunakan Support Vector Machine dan Naive Bayes. *Jurnal Komtika (Komputasi dan Informatika)*, 5(1), 19-25. DOI: <https://doi.org/10.31603/komtika.v5i1.5185>
- [24] Muzaki, A., & Witanti, A. (2021). Sentiment analysis of the community in the twitter to the 2020 election in pandemic covid-19 by method naïve bayes classifier. *Jurnal Teknik Informatika (Jutif)*, 2(2), 101-107. DOI: <https://doi.org/10.20884/1.jutif.2021.2.2.51>
- [25] Fitri, E. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. *Jurnal Transformatika*, 18(1), 71-80. DOI: <http://dx.doi.org/10.26623/transformatika.v18i1.2317>