The Indonesian Journal of Computer Science
www.ijcs.net
Volume 13, Issue 4, August 2024
https://doi.org/10.33022/ijcs.v13i4.4136

Analisis Sentimen Mengenai *Childfree* Menggunakan Metode *Naïve Bayes*

Yeni Safitri[1], Rakhmat Kurniawan[2], Suhardi[3]
yenisafitri0411@gmail.com[1], rakhmat.kr@uinsu.ac.id[2], suhardi@uinsu.ac.id[3]
[1,2,3]Universitas Islam Negeri Sumatera Utara

Informasi Artikel

Diterima : 14 Jun 2024
Direvisi : 1 Jul 2024
Disetujui : 25 Jul 2024

Kata Kunci

analisis sentimen, *childfree*, *naïve bayes*, *twitter*

Abstrak

Kemunculan isu mengenai *childfree* ini menjadi *trending topic* di *Twitter* sejak awal tahun 2020 hingga saat ini, yang banyak melahirkan opini-opini positif dan negatif dari berbagai kalangan khususnya pada media sosial *Twitter*. Penelitian analisis sentimen ini bertujuan untuk mengetahui tanggapan yang diberikan mengenai *childfree* berupa opini positif, netral atau negatif dengan mengumpulkan data *Twitter*. Jumlah *dataset* yang dipakai sebanyak 700 data dengan pembagian sebanyak 630 data latih dan 70 data data uji. Penelitian ini menggunakan klasifikasi Algoritma *Naïve Bayes* dan matriks konfusi sebagai evaluasi kinerja dari sistem yang dibangun. Hasil pengujian menunjukkan nilai akurasi sebesar 64.29%, presisi sebesar 68.25%, *recall* sebesar 64.29% dan *fi-score* sebesar 55.69%.

Keywords

sentiment analysis, childfree, naïve bayes, twitter

Abstract

The emergence of the issue of childfree has become a trending topic on Twitter since the beginning of 2020 until now, which has given rise to many positive and negative opinions from various groups, especially on Twitter social media. This sentiment analysis research aims to determine the responses given regarding childfree in the form of positive, neutral or negative opinions by collecting Twitter data. The number of datasets used is 700 data, divided into 630 training data and 70 test data. This research uses the Naïve Bayes algorithm classification and confusion matrix as a performance evaluation of the system being built. The test results show an accuracy value of 64.29%, precision of 68.25%, recall of 64.29% and fi-score of 55.69%.

A. Pendahuluan

Sebagai negara yang mempunyai total laju kelahiran yang cepat, anak mempunyai peran krusial di kehidupan masyarakat sebab dianggap mampu memberikan berbagai manfaat. Belakangan ini fenomena *childfree* menjadi perbincangan hangat di media sosial *Twitter* yang banyak menimbulkan *pro-kontra* di masyarakat. Istilah *childfree* sendiri merupakan sebuah keputusan yang diambil oleh pasangan suami istri untuk tidak mempunyai anak selama pernikahannya [1]. Penyebutan *childfree* juga tergolong baru ditelinga orang Indonesia, sehingga penyebutan ini tidak mempunyai susunan kata yang dapat diartikan ke bahasa baku yang tidak menyinggung perasaan seseorang. Kemunculan isu mengenai *childfree* ini menjadi *trending topic* di *Twitter* sejak awal tahun 2020 hingga saat ini, yang banyak melahirkan opini positif, netral dan negatif dari berbagai kalangan khususnya pada media sosial *Twitter*. Berdasarkan masalah diatas, peneliti melakukan analisis klasifikasi di media sosial *Twitter* untuk mengetahui opini positif, netral atau negatif mengenai *childfree* dengan algoritma *Naïve Bayes*.

Menghadapi jumlah data yang besar menjadikan orang tidak bisa membaca dan menganalisis data secara manual. Sehingga untuk mengetahui sentimen yang diberikan lebih ke arah *pro* atau *kontra*, dibuat sebuah sistem yang dilengkapi oleh algoritma yang dapat melakukan analisis sentimen dengan mengumpulkan data *Twitter* untuk mengetahui respon masyarakat terhadap *childfree* apakah cenderung bersifat positif, netral atau negatif [2].

Analisis sentimen ialah salah satu bagian *natural language* yang mengolah kata untuk melakukan pelacakan terhadap kondisi masyarakat mengenai suatu produk atau topik tertentu berupa opini (sentimen) seseorang pada kelas sentimen positif, netral atau negatif [3]. Sentimen positif menyatakan opini yang bersifat mendukung terhadap keputusan untuk *childfree*, sedangkan sentimen negatif menyatakan opini yang bersifat penolakan terhadap pilihan tersebut. Dalam melakukan analisis sentimen diperlukan metode yang mendukung yaitu algoritma *Naïve Bayes*. Metode ini merupakan algoritma klasifikasi sederhana yang dipakai untuk mencari nilai probabilitas tertinggi dalam mengklasifikasikan data uji ke dalam kelas yang paling sesuai [4]. Keunggulan dari *Naïve Bayes* yaitu hanya membutuhkan data latih dalam jumlah yang kecil untuk menentukan prediksi dalam proses klasifikasi dan dianggap cukup efektif untuk mendapatkan akurasi yang tinggi [5].

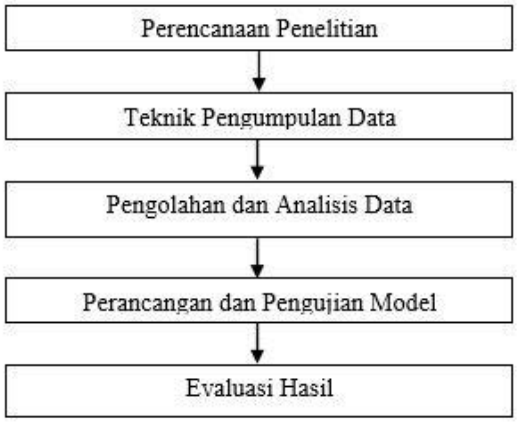
Algoritma *Naïve Bayes* juga sering dipakai oleh para peneliti terdahulu dalam melakukan proses klasifikasi. Penelitian berjudul Analisis Sentimen dan Klasifikasi *Tweet* Berbahasa Indonesia Terhadap Transportasi Umum MRT Jakarta Menggunakan *Naïve Bayes Classifier*[2] bertujuan untuk pengambilan keputusan guna meningkatkan layanan serta fasilitas bagi penumpang MRT dari pengelola. Hasil akurasi yang didapat yaitu 95.88% dengan 70% *precision* positif dan 30% *precision* negatif. Penelitian selanjutnya yaitu Analisis Sentimen Pada *Review* Pengguna *E-Commerce* Menggunakan Algoritma *Naïve Bayes* [6] mengumpulkan data sebanyak 600 data *tweet review* pada *platform e-commerce* Shopee memakai API Shopee. Untuk hasil akurasi didapatkan 99.5%, *precision* sebesar 99.49%, dan *recall* sebesar 100%. Penelitian lainnya berjudul Analisis Sentimen Kebijakan Pembelajaran Tatap Muka Menggunakan *Support Vector Machine* dan *Naïve Bayes*

[7] menghasilkan nilai akurasi *Support Vector Machine* sebesar 88.09% dan *Naïve Bayes* sebesar 75.92%. Output dari analisis ini menyatakan sentimen positif sebanyak 197 dan sentimen negatif sejumlah 3 dari 200 data uji yang digunakan. Jadi, disimpulkan bahwa algoritma *Naïve Bayes* dalam penelitian ini cukup relevan meskipun akurasi yang didapat belum 100%.

Penelitian ini berfokus pada model klasifikasi untuk analisis sentimen berupa informasi tentang sentimen positif, netral dan negatif mengenai *childfree* dari media sosial *Twitter*. Penelitian ini dilakukan untuk mengetahui kinerja dari *Naïve Bayes* dalam melakukan klasifikasi dan mengetahui persentase nilai akurasi, *precision, recall*, dan *f1-score* yang dijalankan oleh sistem menggunakan metode *Naïve Bayes* yang nantinya dapat memberi manfaat tentang penerapan metode *Naïve Bayes* mengenai analisis sentimen serta sebagai referensi tambahan bagi peneliti selanjutnya.

B. Metode Penelitian

Dalam melakukan penelitian analisis sentimen dibutuhkan beberapa tahapan penelitian seperti pada Gambar 1 dibawah ini:



Gambar 1. Tahapan Penelitian

1. Perencanaan Penelitian

Berdasarkan gambar di atas, yang menjadi identifikasi masalah adalah analisis sentimen mengenai *childfree* di *Twitter* yang banyak melahirkan opini-opini positif, netral dan negatif dari berbagai kalangan dengan menerapkan metode *Naïve Bayes*.

2. Teknik Pengumpulan Data

Data yang dipakai pada penelitian ini diperoleh dengan menggunakan teknik *crawling* pada *Twitter* menggunakan kata kunci "*childfree*". Proses *crawling* dilakukan menggunakan *python* dengan bantuan *framework node.js* menggunakan *library tweet-harvest* untuk terhubung ke *Twitter* secara langsung dengan memanfaatkan *API*. Dalam penelitian ini data yang diambil mencakup pendapat atau tanggapan masyarakat mengenai *childfree*. Dari hasil *crawling* tersebut didapat sebanyak 700 data *tweet* dengan pembagian sebanyak 90% sebagai data latih dan 10% sebagai data uji.

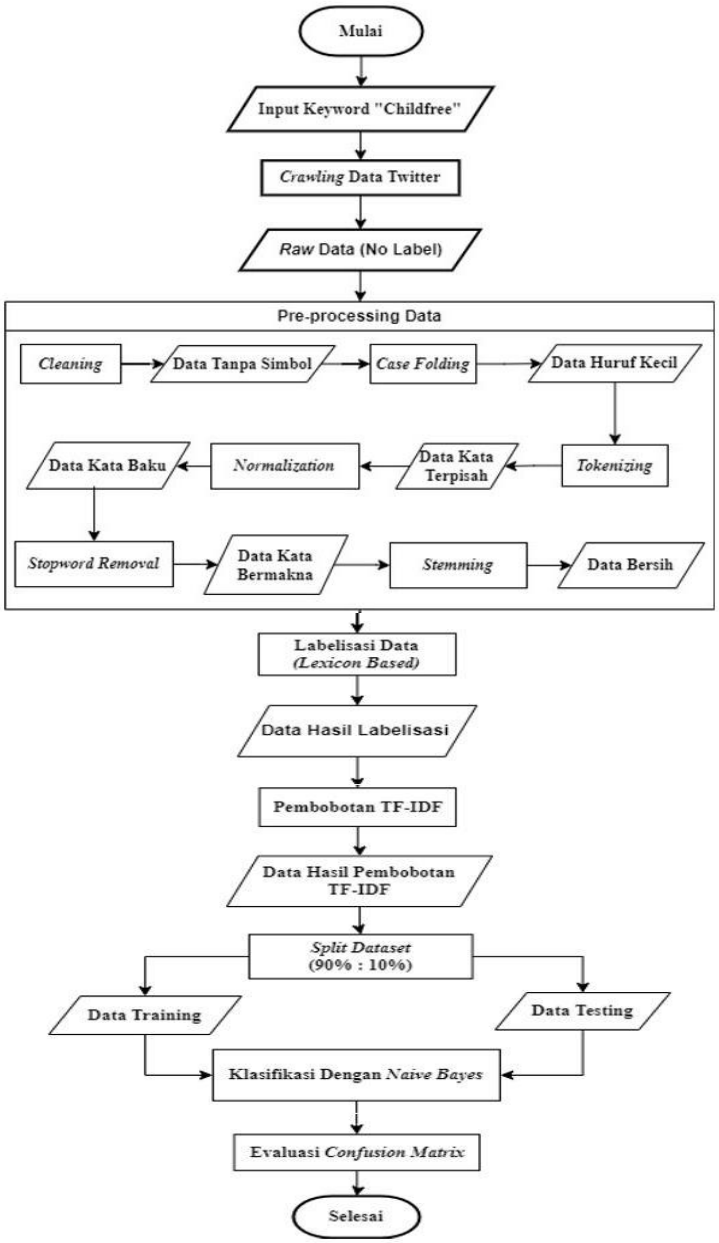
3. Pengolahan dan Analisis Data

Data yang telah didapatkan dari hasil *crawling* selanjutnya diberikan label (kelas) dengan melakukan perhitungan skor sentimen menggunakan kamus

leksikon (*lexicon-based*) yang sudah tersedia. Hasil dari label tersebut berupa kelas positif, kelas netral atau kelas negatif. Data yang telah diberikan label kemudian melewati proses *preprocessing,* lalu masuk ke proses labelisasi data, selanjutnya melakukan pembobotan kata (*term*), kemudian melakukan *splitdata* menjadi data latih dan data uji, lalu data akan siap untuk masuk ke proses klasifikasi. Algoritma *Naïve Bayes* dalam penelitian ini dipakai untuk mengetahui evaluasi kinerja (*confusion matrix*) dari sistem dengan menunjukkan tingkat akurasi dari algoritma tersebut.

4. Perancangan dan Pengujian Model

Rancang sistem dalam penelitian ini dijalankan sesuai dengan kebutuhan serta untuk menyajikan penjelasan yang tepat dan rancang bangunan yang sempurna, seperti gambar dibawah ini:



Gambar 2. Flowchart Sistem Klasifikasi

a. *Preprocessing data*, adalah proses awal pengolahan data yang tujuannya mengurangi bagian yang tidak diinginkan (*noise*) pada data agar data lebih mudah dibaca saat masuk ke dalam proses *mining* [8]. *Preprocessing* merupakan tahap seleksi pemeriksaan pada *text* untuk dilakukan pembersihan, memperbaiki kesalahan yang terdapat dalam *text* serta menyederhanakan *text* agar dapat diproses selanjutnya. Tahapan *preprocessing* adalah sebagai berikut:
   1. *Cleaning,* ialah proses menghapus karakter yang tidak dibutuhkan seperti *username*, *url*, tagar, *mention*, *retweet*, *emoticon*, tanda baca serta simbol *numeric* [9].
   2. *Case folding,* ialah tahapan mengubah seluruh *text* menjadi huruf kecil [10].
   3. *Tokenizing*, merupakan tahapan memotong kata dalam suatu kalimat, paragraf ataupun halaman menjadi kata yang memiliki nilai [9].
   4. *Normalization*, merupakan tahapan memperbaiki kata atau istilah yang tidak benar (*typo*) seperti singkatan dan kata tidak baku agar data bisa dijalankan [11].
   5. *Stopword removal*, ialah tahap menghapus kata yang tidak mempunyai makna seperti kata ganti, kata hubung, dan lainnya [9].
   6. *Stemming*, merupakan tahapan menghilangkan semua kata berimbuhan menjadi kata dasar [12].

b. Labelisasi data, merupakan proses menentukan label (kelas) kedalam kelas positif, netral, atau negatif. Pada tahap ini label (kelas) ditentukan dengan menghitung nilai skor sentimennya yang mengacu pada kamus *lexicon based* yang dipakai [13].

c. Pembobotan Kata *TF-IDF*, berfungsi sebagai penghitung nilai frekuensi sebuah kata (*term*) yang terdapat di dalam dokumen. Algoritma ini dipakai untuk menghitung bobot dengan menyatukan dua pola yakni banyaknya nilai kemunculan jumlah kata dalam dokumen tertentu dan banyaknya nilai kemunculan jumlah kata dari dokumen yang mengandung kata tersebut [14].

$$W_{dt} = TF_{dt} \times IDF \quad (1)$$

$$\text{Norm } W_{dt} = \frac{W_{dt}}{\sqrt{\sum_{i-1}^{n} x_i^2}} \quad (2)$$

d. Klasifikasi *Naïve Bayes*, ialah klasifikasi menggunakan pendekatan probabilistik dan statistik [15]. Keunggulan metode ini yaitu dapat menggunakan sedikit data latih untuk proses klasifikasi.

$$P(H|X) = \frac{P(H)P(X|H)}{P(X)} \quad (3)$$

e. Evaluasi, merupakan tahap melakukan pengujian dari sistem yang dibangun menggunakan *confusion matrix* sebagai bahan evaluasi hasil. *Confusion matrix* menyajikan tabel matrix yang terdiri dari kelas positif dan kelas negatif [16].

Tabel 1. Confusion Matrix

Actual Label	***Predict Label***	
	Positive (P)	*Negative (N)*
Positive (P)	*True Positive (TP)*	*False Positive (FP)*
Negative (N)	*False Negative (FN)*	*True Negative (TN)*

C. Hasil dan Pembahasan

1. Analisis data

Data yang dipakai pada penelitian ini mencakup tanggapan umum di media sosial Twitter yang berkaitan dengan *childfree*. Data tersebut didapatkan dari hasil *crawling* dengan menggunakan kata kunci #*childfree*. Data yang berhasil dikumpulkan sebanyak 700 data *tweet* yang dibagi menjadi 90% sebagai data *training* dan 10% sebagai data *testing*. Sehingga sebanyak 630 data digunakan sebagai data pelatihan dan 70 data digunakan sebagai data pengujian.

Tabel 2. Sampel Data Training

Data *Training*
kalo childfree. nanti tumbalnya diri sendiri. orang tua atau bisa jadi saudara
@kochengfs harusnya emang bener. childfree aja"""""
saya abis belajar maternitas malah makin mau childfree
@urfavpartnersx bener. Di indonesia yang mau childfree malah dihujat. Miris
Alasan Childfree wajib hukumnya

Tabel 3. Sampel Data Testing

Data Testing
Childfree dan Ancaman Musnahnya Peradaban di Masa Depan https://t.co/ZOprNammfl

2. *Preprocessing* Data

Data hasil *crawling* selanjutnya harus melewati proses *preprocessing* karena masih banyaknya data yang memiliki elemen yang tidak diperlukan (*noise*) dalam proses analisis sentimen.

Tabel 4. Hasil Preprocessing

Sentimen Awal	**Sentimen Hasil Preprocessing**
Data Training	
kalo childfree. nanti tumbalnya diri sendiri. orang tua atau bisa jadi saudara	kalo childfree tumbal orang tua saudara
@kochengfs harusnya emang bener. childfree aja"""""	emang bener childfree aja
saya abis belajar maternitas malah makin mau childfree	abis ajar maternitas childfree
@urfavpartnersx bener. Di indonesia yang mau childfree malah dihujat. Miris	bener indonesia childfree hujat miris
Alasan Childfree wajib hukumnya	alas childfree wajib hukum
Data Testing	
Childfree dan Ancaman Musnahnya Peradaban di Masa Depan https://t.co/ZOprNammfl	childfree ancam musnah adab

3. Pembobotan *TFI-DF*

Dari data hasil *preprocessing* pada tabel 4, dapat diperoleh nilai kemunculan jumlah kata (*term*). Berikut proses perhitungan *TF-IDF*:

Tabel 5. Pembobotan TF-IDF

Term	***TF***					***DF***
	D1	**D2**	**D3**	**D4**	**D5**	
abis	0	0	1	0	0	1
aja	0	1	0	0	0	1
ajar	0	0	1	0	0	1
alas	0	0	0	0	1	1
bener	0	1	0	1	0	2
childfree	1	1	1	1	1	5
emang	0	1	0	0	0	1
hujat	0	0	0	1	0	1
hukum	0	0	0	0	1	1
indonesia	0	0	0	1	0	1
kalo	1	0	0	0	0	1
maternitas	0	0	1	0	0	1
miris	0	0	0	1	0	1
orang	1	0	0	0	0	1
saudara	1	0	0	0	0	1
tua	1	0	0	0	0	1
tumbal	1	0	0	0	0	1
wajib	0	0	0	0	1	1

Selanjutnya menghitung nilai *IDF* untuk *term* pertama kata “abis” dimana jumlah *d* = 5 dan *df* =1.

$$\text{IDF} = \ln \frac{5+1}{1+1} + 1 = 2{,}098612 \quad (4)$$

Selanjutnya tahap perhitungan nilai *TF-IDF* untuk *term* pertama pada kata “abis” dimana nilai *tf* = 1 dan nilai *idf* =2,098612

$$W_{dt} = 1 \times 2{,}098612 = 2{,}098612 \quad (5)$$

Tabel 6. Hasil TF-IDF (W_{dt})

Term	***TF-IDF* (*Wdt*)**				
	D1	**D2**	**D3**	**D4**	**D5**
abis	0	0	2,09861	0	0
aja	0	2,09861	0	0	0
ajar	0	0	2,09861	0	0
alas	0	0	0	0	2,09861
bener	0	1,69315	0	1,69315	0
childfree	1	1	1	1	1
emang	0	2,09861	0	0	0

Term	*TF-IDF* (*Wdt*)				
	D1	**D2**	**D3**	**D4**	**D5**
hujat	0	0	0	2,09861	0
hukum	0	0	0	0	2,09861
indonesia	0	0	0	2,09861	0
kalo	2,09861	0	0	0	0
maternitas	0	0	2,09861	0	0
miris	0	0	0	2,09861	0
orang	2,09861	0	0	0	0
saudara	2,09861	0	0	0	0
tua	2,09861	0	0	0	0
tumbal	2,09861	0	0	0	0
wajib	0	0	0	0	2,09861

Nilai *TF-IDF* W_{dt} dinormalisasikan agar nilai bobot berada pada *range* yang sama dengan rumus:

$$\text{Norm } W_{dt} = \frac{W_{dt}}{\sqrt{\sum_{i-1}^{n} x_i^2}} \quad (6)$$

Tabel 7. Nilai Normalisasi TF-IDF (W_{dt})

Term	**Normalisasi *TF-IDF* (*Wdt*)**				
	D1	**D2**	**D3**	**D4**	**D5**
abis	0	0	0,55667	0	0
aja	0	0,58946	0	0	0
ajar	0	0	0,55667	0	0
alas	0	0	0	0	0,55667
bener	0	0,47558	0	0,40969	0
childfree	0,20842	0,28088	0,26526	0,24197	0,26526
emang	0	0,58946	0	0	0
hujat	0	0	0	0,50781	0
hukum	0	0	0	0	0,55667
indonesia	0	0	0	0,50781	0
kalo	0,43739	0	0	0	0
maternitas	0	0	0,55667	0	0
miris	0	0	0	0,50781	0
orang	0,43739	0	0	0	0

Term	Normalisasi *TF-IDF* (*Wdt*)				
	D1	**D2**	**D3**	**D4**	**D5**
saudara	0,43739	0	0	0	0
tua	0,43739	0	0	0	0
tumbal	0,43739	0	0	0	0
wajib	0	0	0	0	0,55667

4. Klasifikasi *Naïve Bayes*

Setelah proses pembobotan selesai, pada tahap ini dilakukan perhitungan klasifikasi terhadap data uji untuk memprediksi *label* pada data uji berdasarkan data latih yang telah diketahui labelnya.

Tabel 8. Data Latih

Dokumen	Data Opini	Label
D1	kalo childfree tumbal orang tua saudara	Negatif
D2	emang bener childfree aja	Positif
D3	abis ajar maternitas childfree	Netral
D4	bener indonesia childfree hujat miris	Negatif
D5	alas childfree wajib hukum	Negatif

Tabel 9. Data Uji

Data *tweet*	Hasil *Preprocessing*	Label
Childfree dan Ancaman Musnahnya Peradaban di Masa Depan https://t.co/ZOprNammfl	childfree ancam musnah adab	?

Berikut tahap perhitungan metode *Naïve Bayes* terhadap data uji.

1) Nilai *prior probability*

a. P(Positif) $= \frac{1}{5} = 0{,}2$

b. P(Negatif) $= \frac{3}{5} = 0{,}6$

c. P(Netral) $= \frac{1}{5} = 0{,}2$

2) Nilai *conditional probability*

a. Positif

P(childfree|Positif) $= \frac{1+1}{6{,}89 + 36{,}27} = 0{,}046$

P(ancam|Positif) $= \frac{0+1}{6{,}89 + 36{,}27} = 0{,}023$

P(musnah|Positif) $= \frac{0+1}{6{,}89+36{,}27} = 0{,}023$

P(adab|Positif) $= \frac{0+1}{6{,}89+36{,}27} = 0{,}023$

b. Negatif

P(childfree|Negatif) $= \frac{3+1}{27{,}78+36{,}27} = 0{,}062$

P(ancam| Negatif) $= \frac{0+1}{27{,}78+36{,}27} = 0{,}016$

P(musnah| Negatif) $= \frac{0+1}{27{,}78+36{,}27} = 0{,}016$

P(adab|Positif) $= \frac{0+1}{27{,}78+36{,}27} = 0{,}016$

c. Netral

P(childfree| Netral) $= \frac{1+1}{7{,}30+36{,}27} = 0{,}046$

P(ancam| Netral) $= \frac{0+1}{7{,}30+36{,}27} = 0{,}023$

P(musnah| Netral) $= \frac{0+1}{7{,}30+36{,}27} = 0{,}023$

P(adab|Positif) $= \frac{0+1}{7{,}30+36{,}27} = 0{,}023$

3) Nilai *posterior probability*

P(Opini|Positif) $= 0{,}046 * 0{,}023^3$
$= 5{,}5968 * 10^{-7}$

P(Opini|Negatif) $= 0{,}062 * 0{,}016^3$
$= 2{,}5395 * 10^{-7}$

P(Opini|Netral) $= 0{,}046 * 0{,}023^3$
$= 5{,}5968 * 10^{-7}$

Berdasarkan perhitungan di atas didapatkan hasil paling tinggi dengan nilai **$2{,}5395*10^{-7}$** dengan label **Negatif**. Sehingga hasil klasifikasi untuk data uji adalah **Negatif**.

5. Evaluasi

Setelah melakukan proses pengujian pada algoritma *Naïve Bayes*, selanjutnya melakukan pengujian model menggunakan 70 data uji dengan menghitung *Confusion Matrix*. Berikut hasil pengujian model menggunakan data pengujian:

Tabel 10. Confusion Matrix

ij		**Kelas Prediksi (*j*)**		
		Positif	**Negatif**	**Netral**
Kelas Aktual (*i*)	**Positif**	1	10	3
	Negatif	0	41	1
	Netral	0	11	3

Dari tabel di atas dapat dihitung nilai *accuracy*, *precision*, *recall*, dan *f1-score* dengan rumus berikut:

$$Accuracy = \frac{1+41+3}{1+10+3+0+41+1+0+11+3} \times 100\% = 64.29\%$$

$$Precision = \frac{45}{41+3+1} \times 100\% = 68.25\%$$

$$Recall = \frac{45}{41+3+1} \times 100\% = 64.29\%$$

$$f1score = \frac{2 \times 68.25 \times 64.29}{68.25 + 64.29} \times 100\% = 55.69\%$$

Untuk melihat nilai *accuracy* secara keseluruhan dari hasil analisis sentimen dapat menggunakan *classification report* berikut:

```
Confusion Matrix:

[[41  1  0]
 [11  3  0]
 [10  3  1]]


Test Accuracy   : 64.29%
Test Recall     : 64.29%
Test Precision  : 68.25%
Test F1-Score   : 55.69%


Classification Report:

              precision    recall  f1-score   support

    Negative       0.66      0.98      0.79        42
     Neutral       0.43      0.21      0.29        14
    Positive       1.00      0.07      0.13        14

    accuracy                           0.64        70
   macro avg       0.70      0.42      0.40        70
weighted avg       0.68      0.64      0.56        70
```

Gambar 3. Classification Report

Dari perhitungan di atas didapatkan nilai akurasi sebesar 64.29%, *precision* 68.25%, *recall* 64.29%, dan *f1-score* sebesar 55.69%. Hal ini menunjukkan bahwa model yang dibangun cukup baik dan memberikan hasil prediksi yang cukup baik juga pada data yang belum diketahui labelnya.

D. Simpulan

Berdasarkan hasil dari penelitian analisis sentimen mengenai *childfree* dengan metode *Naïve Bayes* untuk menganalisis klasifikasi teks komentar atau opini masyarakat, diperoleh sebanyak 179 data kelas positif, 216 data kelas netral dan 305 data kelas negatif dari 700 data yang dipakai dengan teknik pelabelan menggunakan *lexicon-based*. Sehingga didapatkan nilai akurasi sebesar 64.29%, *precision* 68.25%, *recall* 64.29%, dan *f1-score* sebesar 55.69%. Kondisi ini menerangkan jika algoritma klasifikasi *Naïve Bayes* bekerja dengan baik pada data yang digunakan. Untuk penelitian selanjutnya diharapkan menggunakan data yang didalamnya terkandung kalimat sarkas. Dapat juga mengembangkan sistem mengenai pembangunan *user interface* yang menyerupai *web*, *mobile*, maupun *desktop* agar memudahkan pengguna lainnya membaca output dari analisis sentimen serta dapat melakukan analisis sentimen yang dapat menjangkau umur yang berkomentar.

E. Referensi

[1] E. Fadhilah, "Childfree dalam Perspektif Islam," *al-Mawarid J. Syari'ah Huk.*, vol. 1, hal. 71–80, 2021, doi: 10.20885/mawarid.vol3.iss2.art1.

[2] D. Ikasari, Y. Fajarwati, dan Widiastuti, "Analisis Sentimen Dan Klasifikasi Tweets Berbahasa Indonesia Terhadap Transportasi Umum Mrt Jakarta Menggunakan Naïve Bayes Classifier," *J. Ilm. Inform. Komput.*, vol. 25, no. 1, hal. 64–75, 2020, doi: 10.35760/ik.2020.v25i1.2427.

[3] K. V. S. Toy, Y. A. Sari, dan I. Cholissodin, "Analisis Sentimen Twitter menggunakan Metode Naive Bayes dengan Relevance Frequency Feature Selection (Studi Kasus: Opini Masyarakat mengenai Kebijakan New Normal)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 11, hal. 5068–5074, 2021, [Daring]. Tersedia pada: http://j-ptiik.ub.ac.id

[4] B. Z. Ramadhan, I. Riza, dan I. Maulana, "Analisis Sentimen Ulasan pada Aplikasi E-Commerce dengan Menggunakan Algoritma Naïve Bayes," *J. Appl. Informatics Comput.*, vol. 6, hal. 220–225, 2022.

[5] A. Harun dan D. P. Ananda, "Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naïve bayes dan Decission Tree," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, hal. 58–63, 2021.

[6] A. H. Hasugian, M. Fakhriza, dan D. Zukhoiriyah, "Analisis Sentimen Pada Review Pengguna E-Commerce Menggunakan Algoritma Naïve Bayes," *Januari*, vol. 6, no. 1, hal. 98–107, 2023, [Daring]. Tersedia pada: https://ojs.trigunadharma.ac.id/index.php/jsk/index

[7] M. S. Hasibuan dan A. Serdano, "Analisis Sentimen Kebijakan Pembelajaran Tatap Muka Menggunakan Support Vector Machine dan Naive Bayes," *JRST (Jurnal Ris. Sains dan Teknol.*, vol. 6, no. 2, hal. 199–204, 2022, doi: 10.30595/jrst.v6i2.15145.

[8] M. N. Rizaldi, Adiwijaya, dan S. Al Faraby, "Klasifikasi Argument Pada Teks dengan Menggunakan Metode Multinomial Logistic Regression Terhadap Kasus Pemindahan Ibu Kota Indonesia di Twitter," *Media Inform. Budidarma*, vol. 4, no. 4, hal. 904–913, 2020, doi: 10.30865/mib.v4i4.2348.

[9] M. Furqan, S. Sriani, dan S. Mayang Sari, "Analisis Sentimen Menggunakan K-Nearest Neighbor Terhadap New Normal Masa Covid-19 Di Indonesia," *Techno.Com*, vol. 21, no. 1, hal. 52–61, 2022, doi: 10.33633/tc.v21i1.5446.

[10] H. Santoso, Armansyah, dan D. Desliani, "Analisis Sentimen Mahasiswa Terkait Pembelajaran Tatap Muka Menggunakan Metode Naive Bayes Classifier," *Techno.Com*, vol. 21, hal. 644–654, 2022, doi: 10.33633/tc.v21i3.6262.

[11] Y. A. N. Jannah dan R. B. Prasetyo, "Analisis Sentimen dan Emosi Publik pada Awal Pandemi COVID-19 Berdasarkan Data Twitter dengan Pendekatan Berbasis Leksikon," *Semin. Nas. Off. Stat. 2022*, hal. 597–607, 2022.

[12] A. Rahman, E. Utami, dan Sudarmawan, "Sentimen Analisis Terhadap Aplikasi pada Google Playstore Menggunakan Algoritma Naïve Bayes dan Algoritma Genetika," *J. Komtika (Komputasi dan Inform.*, vol. 5, no. 1, hal. 60–71, 2021.

[13] D. Zukhoiriyah, "Analisis Sentimen Pada Review Pengguna E-Commerce Menggunakan Algoritma Naive Bayes (Studi Kasus: Shopee)," UIN Sumatera Utara, 2022.

[14] I. Di Estika, I. Darmawan, dan O. N. Pratiwi, "Analisis Sentimen Ulasan Pengguna Untuk Peningkatan Layanan Menggunakan Algoritma Naive Bayes

(Studi kasus: Bukalapak)," *e-Proceeding Eng.*, vol. 8, no. 2, hal. 2735–2745, 2021.

[15] D. P. Utomo dan Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *J. Media Inform. Budidarma*, vol. 4, no. April, hal. 437–444, 2020, doi: 10.30865/mib.v4i2.2080.

[16] Suherman, M. Purnamasari, dan F. D. Hastuti, "Klasifikasi Siswa Berdasarkan Mata Pelajaran Lintas Minat Menggunakan Metode Decision Tree C4.5," *JSiI (Jurnal Sist. Informasi)*, vol. 8, no. 2, hal. 141–149, 2021.