

## ANALISIS SENTIMEN TERHADAP PEMBERHENTIAN TV ANALOG PADA TWITTER MENGGUNAKAN ALGORITMA NAIVE BAYES

Andika Prasetyo<sup>1</sup>, Taufik Ridwan<sup>2</sup>

<sup>1,2</sup> Prodi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang  
Jl. HS. Ronggowaluyo, Telukjambe Timur, Karawang  
(0267) 641177

E-mail: [andika.prasetyo19004@student.unsika.ac.id](mailto:andika.prasetyo19004@student.unsika.ac.id)<sup>1</sup>, [taufik.ridwan@cs.unsika.ac.id](mailto:taufik.ridwan@cs.unsika.ac.id)<sup>2</sup>

### ABSTRACT

*Technological developments make changes that occur in television broadcasts which were originally analog to digital. This rule makes the subject of discussion among the general public, one of which is Twitter users. This raises the opinions of Twitter users pros and cons. Regarding this, there is a need for sentiment analysis of analog TV stops using the Naive Bayes algorithm. The analysis was carried out to find out the opinions of Twitter users regarding the discontinuation of analog tv. The methodology applied uses the AI project cycle which ends with the evaluation stage. Data retrieved with the keyword "tv analog" in the period 1 January 2022 to December 30, 2022 with a maximum limit of 100 data. The data obtained is given a sentiment label with two sentiments, namely Negative and Positive. In this case, it produces negative sentiment data of 70 and positive sentiment of 30. In classifying data using the Naive Bayes algorithm to obtain a classification model using the RapidMiner Studio 10.0 tool. Testing the best model is done by means of five tests with the distribution of training data and test data 50:50 to 90:10. The best model is found in the five tests with 90% accuracy. The best model performance gets 90% accuracy, 75% precision, and 100% recall where these results are good.*

**Keywords:** *AI Project Cycle, Analog TV, Naive Bayes, Sentiment, Twitter.*

### ABSTRAK

Perkembangan teknologi menjadikan perubahan yang terjadi pada siaran televisi yang semulanya analog menjadi digital. Aturan tersebut menjadikan bahan perbincangan pada kalangan umum salah satunya pengguna Twitter. Hal itu menimbulkan opini pengguna Twitter pro-kontra. Mengenai hal ini perlu adanya analisis sentimen terhadap pemberhentian tv analog menggunakan algoritma *Naive Bayes*. Analisis dilakukan untuk mengetahui opini pengguna Twitter terhadap pemberhentian tv analog. Metodologi yang diterapkan menggunakan *AI project cycle* yang diakhiri dengan tahap *evaluation*. Data yang diambil dengan kata kunci "tvanalog" dalam kurun waktu 1 Januari 2022 s.d. 30 Desember 2022 dengan batas maksimal 100 data. Data yang didapatkan diberikan label sentimen dengan dua sentimen yaitu Negatif dan Positif. Dalam hal tersebut menghasilkan data sentimen Negatif berjumlah 70 dan sentimen Positif berjumlah 30. Dalam pengklasifikasian data menggunakan algoritma *Naive Bayes* untuk mendapatkan model klasifikasi dengan menggunakan alat bantu RapidMiner Studio 10.0. Pengujian model terbaik dilakukan dengan cara lima pengujian dengan pembagian data latih dan data uji 50:50 sampai dengan 90:10. Model terbaik terdapat pada pengujian lima dengan hasil akurasi 90%. Performa model terbaik mendapatkan hasil akurasi 90%, presisi 75%, dan *recall* 100% dimana hasil tersebut baik.

**Kata kunci:** *AI Project Cycle, Naive Bayes, Sentimen, TV Analog, Twitter.*

### 1. PENDAHULUAN

Pada perkembangan teknologi masa sekarang setiap perubahan terjadi dimana siaran Televisi (TV) pun mengalami perubahan yang semulanya analog kini bermigrasi ke digital sebagaimana yang tertuang pada UU No 11 Tahun 2020 tentang Cipta Kerja Pasal 60A ayat (1) yang bertuliskan "Penyelenggaraan penyiaran dilaksanakan dengan mengikuti perkembangan teknologi, termasuk migrasi penyiaran dari teknologi analog ke teknologi digital." (Presiden Republik Indonesia, 2020). Hal ini membuat masyarakat harus mengikuti aturan pemerintah tersebut untuk bermigrasi ke TV digital. Bilamana TV tidak mendukung siaran digital dapat

menambahkan *Set Top Box* untuk mendapatkan siaran digital.

Dilihat dari aspek pemberhentian TV analog masyarakat memberikan opini ataupun pendapatnya melalui berbagai media, salah satunya media sosial Twitter. Twitter adalah media sosial yang populer di masyarakat (Pintoko & L., 2018). Pendapat yang diberikan terhadap Program Analog *Switch Off* (ASO) atau pemberhentian TV analog beragam, seperti mendukung dan kurang mendukung. Dalam hal ini bilamana diteliti lebih lanjut terhadap twit "tvanalog" akan mendapatkan sentimen yang berbeda-beda pada setiap opini yang ada. Hal tersebut bila dikumpulkan akan menghasilkan informasi. Tugas mendasar analisis sentimen yaitu

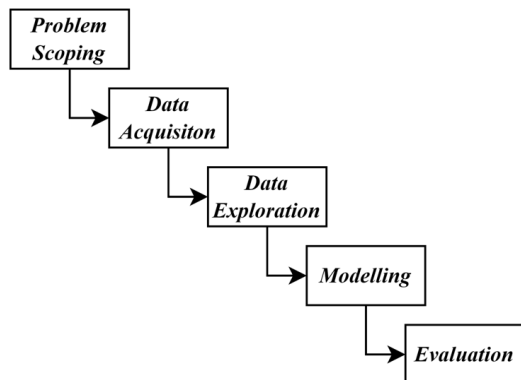
mengklasifikasikan polaritas kalimat yang ada di dokumen. Polaritas mengacu pada arti sebuah teks yang ada di dalam dokumen memiliki sentimen negatif atau positif (Nugroho et al., 2016).

Dalam hal tersebut perlu adanya algoritma *text mining* yang diperlukan, dalam penelitian ini menggunakan algoritma *Naive Bayes*. Sebagaimana penelitian yang ada mengenai implementasi algoritma *Naive Bayes* terhadap analisis sentimen opini film pada Twitter menghasilkan akurasi 90%, *precision* 92%, *recall* 90% dan *f-measure* 90% (Ratnawati, 2018). Dalam penelitian tersebut algoritma *Naive Bayes* layak untuk digunakan pada penelitian ini. Dalam operasinya akan dibantu menggunakan *tools* RapidMiner Studio 10.0.

Tujuan penelitian ini untuk mengetahui dan menghasilkan analisis sentimen TV analog menggunakan algoritma *Naive Bayes* dan hasil performa algoritma yang digunakan. Penelitian ini juga bisa dijadikan referensi untuk penelitian yang serupa.

## 2. METODE

Penelitian ini akan menggunakan metodologi *AI project cycle* dalam proses pengimplementasian dengan model *Naive Bayes*. Tahapan yang dilakukan terdapat pada Gambar 1.



Gambar 1. Tahapan *AI Project Cycle*

### 2.1 *AI Project Cycle*

*AI project cycle* merupakan himpunan rangkaian untuk memecahkan suatu permasalahan dalam suatu penelitian menggunakan metode dan kaidah ilmiah yang kredibel dan menarik kesimpulan mengenai hal itu. Dalam hal ini, metode *AI project cycle* memiliki beberapa tahapan diantaranya *problem scoping*, *data acquisition*, *data exploration*, *modelling*, dan *evaluation* (Halim et al., 2022).

#### 2.1.1 *Problem Scoping*

*Problem scoping* merupakan rangkaian pertama dari *AI project cycle* dengan mengidentifikasi atau menentukan batasan masalah yang bertujuan agar objek penelitian dapat lebih jelas serta terarah. Adapun dalam *problem scoping*,

terdapat metode 4W yang terdiri dari *who*, *where*, *what*, dan *why* yang dimaksudkan untuk mempermudah proses *problem scoping* (Azimah et al., 2022).

#### 2.1.2 *Data Acquisition*

*Data acquisition* adalah rangkaian kedua setelah selesai mengidentifikasi masalah, yang kemudian peneliti beralih pada apa yang dibutuhkan terkait penyelesaian terkait data (Widodo et al., 2022). Pada tahap ini penelitian menggunakan data yang diambil melalui Twitter terkait tv analog dalam kurun waktu 1 Januari 2022 s.d. 30 Desember 2022 dengan batas maksimal 100 data dengan menggunakan alat bantu *snsrape*. Tahap ini menggunakan alat bantu bernama Google Colab.

#### 2.1.3 *Data Exploration*

Diperkenalkan oleh John Tukey, *data exploration* merupakan rangkaian penindaklanjutan data yang didapat dari *data acquisition*, pada tahap ini data diatur dan dikumpulkan secara seragam dengan pemahaman karakteristik yang didapatkan dengan *preprocessing* terkait data yang diperoleh tersebut (Halim et al., 2022). Pada tahap ini akan menampilkan data yang telah didapatkan dan memberikan sentimen terhadap data yang didapatkan lalu ditampilkan. Setelah itu akan dilakukan *preprocessing* terhadap data yang didapatkan untuk diproses pada tahap selanjutnya. Tahap ini menggunakan alat bantu Microsoft Excel dan RapidMiner Studio 10.0

#### 2.1.4 *Modelling*

Setelah melewati *data exploration*, tahap selanjutnya adalah *modelling* yaitu pengembangan model yang mencakup proses pemilihan algoritma dan *training data*. Proses pembuatan algoritma dalam bahasa pemrograman dimaksudkan sebagai metode pembelajaran mesin (*training phase*) yang digunakan untuk menemukan pola (Hozairi et al., 2021). Mengenai hal itu, model klasifikasi yang digunakan dalam penelitian ini menggunakan *Naive Bayes*. Pada tahap ini dilakukan pemisahan data latih dan data uji dengan lima percobaan 50:50 sampai dengan 90:10. Percobaan akurasi terbaik akan digunakan dalam menentukan evaluasi. Tahap ini menggunakan alat bantu RapidMiner Studio 10.0.

*Naive Bayes* adalah metode klasifikasi yang efisien dan sangat sederhana. Klasifikasi *Bayesian* berdasarkan teorema *Bayes* memiliki kemampuan untuk mengklasifikasikan dengan akurasi dan kecepatan tinggi ketika diterapkan pada *database* yang berisi data dalam jumlah besar. Berikut adalah bentuk umum teorema *Bayes* (Christianto et al., 2020):

$$P(H|X) = \frac{p(X|H) \times p(H)}{p(X)} \quad (1)$$

Keterangan:

$P(H|X)$  = Probabilitas hipotesis H berdasarkan kondisi X

$P(H)$  = Probabilitas hipotesis H

$P(X|H)$  = Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$  = Probabilitas dari X

Proses selanjutnya memerlukan penunjuk dalam menentukan kelas yang cocok pada sampel yang dianalisis. Maka dari itu, persamaan (1) disesuaikan (Christianto et al., 2020):

$$P(C|F_1 \dots F_n) = \frac{P(C) \times p(F_1 \dots F_n|C)}{p(F_1 \dots F_n)} \quad (2)$$

Keterangan:

$P(C|F_1 \dots F_n)$  = Nilai probabilitas kelas C dari karakteristik  $F_1 \dots F_n$

$P(C)$  = Probabilitas kelas C

$F_1 \dots F_n$  = Karakteristik petunjuk untuk melakukan klasifikasi

$p(F_1 \dots F_n|C)$  = Nilai probabilitas karakteristik  $F_1 \dots F_n$  pada kelas C

$p(F_1 \dots F_n)$  = Probabilitas karakteristik  $F_1 \dots F_n$

Suatu dataset kemungkinan adanya nilai nol pada model probabilitas. Nilai nol tersebut membuat model *Naive Bayes* tidak dapat melakukan klasifikasi, oleh karena itu diperlukan metode yang dapat menyelesaikan masalah tersebut. *Add one smoothing* atau *Laplacian Smoothing* dapat menjadi solusi untuk menghindari angka nol, karena dalam perhitungan parameter  $n_k$  ditambah dengan satu (Christianto et al., 2020).

$$P(F_k|C_i) = \frac{n_k+1}{n+|Vocabulary|} \quad (3)$$

Keterangan:

$P(F_k|C_i)$  = Nilai probabilitas  $F_k$  pada kelas  $C_i$

$n_k$  = Jumlah sampel data atribut  $n_k$

$n$  = Jumlah sampel data dalam kelas dari atribut  $n_k$

$|Vocabulary|$  = Jumlah sampel data

### 2.1.5 Evaluation

*Evaluation* merupakan rangkaian lanjutan dari *modelling*. Pada tahap ini, model diuji dan dianalisis mengenai permasalahan dan sebab-akibat dalam hasil dan penggunaan model dan data set. Uji analisis dilakukan dalam banyak cara termasuk di dalamnya evaluasi dilakukan dengan memperhatikan akurasi dari data tersebut setelah *training* apakah sudah baik atau belum. (Mufidah et al., 2022). Pada tahap ini akan digunakan *confusion matrix* dalam mengevaluasi model yang terbaik. Tahap ini menggunakan alat bantu RapidMiner Studio 10.0.

*Confusion matrix* merupakan metode yang tepat digunakan untuk melakukan perhitungan

akurasi pada *text mining*. *Confusion matrix* diilustrasikan dengan tabel yang menunjukkan jumlah data uji yang diklasifikasikan dengan benar dan data uji yang salah diklasifikasikan. Tabel dari *confusion matrix* sebagai berikut (Christianto et al., 2020):

**Tabel 1.** *Confusion Matrix*

<i>Confusion Matrix</i>		Prediksi	
		Positif	Negatif
Aktual	Positif	TP ( <i>True Positif</i> )	FN ( <i>False Negatif</i> )
	Negatif	FP ( <i>False Positif</i> )	TN ( <i>True Negatif</i> )

Berdasarkan nilai dari *confusion matrix* tersebut, dapat menghasilkan nilai akurasi, presisi, dan *recall* dengan persamaan sebagai berikut (Fikri et al., 2020):

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (4)$$

$$\text{Presisi} = \frac{TP}{TP+FP} \times 100\% \quad (5)$$

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\% \quad (6)$$

## 3. PEMBAHASAN

### 3.1 Problem Scoping

*Problem scoping* ditentukan dengan 4 langkah, yaitu:

1. *Who* (siapa yang akan mengetahui informasi tersebut): hasil penelitian ini akan memberikan informasi bagi pembaca mengenai analisis sentimen pemberhentian tv analog dan performa algoritma yang digunakan.
2. *What* (apa masalah tersebut): sejak diberlakukannya pemberhentian tv analog opini masyarakat pro-kontra, mengenai hal itu maka diperlukannya analisis sentimen terhadap pemberhentian tv analog. Dari hasil analisis sentimen digunakanlah algoritma *Naive Bayes* untuk pengklasifikasian sentimen pada opini tersebut.
3. *Where* (dimana masalah tersebut): media sosial salah satu tempat setiap orang mencurahkan opininya karena disitu dapat bebas berpendapat. Media sosial yang digunakan adalah Twitter.
4. *Why* (mengapa masalah dibuat): untuk mengetahui hasil mengenai analisis sentimen pemberhentian tv analog dan mengetahui performa algoritma *Naive Bayes* pada analisis sentimen pemberhentian tv analog.

### 3.2 Data Acquisition

Untuk pengumpulan data menggunakan cara *scraping* pada Twitter. *Scraping* dilakukan menggunakan Google Colab dengan bahasa yang digunakan ialah Python dan menggunakan *library*

bernama *snsrape*. Data yang diambil dengan kata kunci “tvanalog” dalam kurun waktu 1 Januari 2022 s.d. 30 Desember 2022 dengan batas maksimal 100 data. Atribut data yang diambil yaitu *user*, waktu, *like* pada *tweet*, sumber *tweet*, dan isi *tweet*. *Source code* dapat dilihat pada Gambar 2.

```
!pip install snsrape
import snsrape.modules.twitter as sntwitter
import pandas as pd

# Membuat daftar untuk menambahkan data tweet
tweets = []

# Menggunakan TwitterSearchScrapper untuk mengikis data dan menambahkan
tweet ke daftar
for i, tweet in enumerate(sntwitter.TwitterSearchScrapper('tvanalog
since:2022-12-01 until:2022-12-30 lang:id').get_items()):
    if i>99:
        break
    tweets.append([tweet.user.username, tweet.date, tweet.likeCount,
tweet.sourceLabel, tweet.content])

# Membuat kerangka data untuk memuat daftar
tweets_df = pd.DataFrame(tweets, columns=["User", "Date Created", "Number
of Likes", "Source of Tweet", "Tweet"])

tweets_df.to_csv('sentimenttvanalog.csv', index=False)
```

Gambar 2. Source Code Pengambilan Data

### 3.3 Data Exploration

Pada proses ini akan menampilkan data lima teratas dan lima terbawah yang dapat dilihat pada Gambar 3. Data tersebut akan dilakukan pelabelan manual untuk mengetahui sentimen pada *tweet* termasuk negatif atau positif.

User	Date Created	Number of Likes	Source of Tweet	Tweet	
0	el_echo	2022-12-29 23:55:33+00:00	2	Twitter for Android	wing2 lagak'e paling teranyaya pas tv analog ...
1	ibkindartan48	2022-12-29 23:51:13+00:00	7	Twitter for Android	Gue masih menggunakan TV analog bukan TV digita...
2	delikinet	2022-12-29 23:47:28+00:00	0	Echobox	Pemadaman siaran TV analog dan dialihkan ke TV...
3	lukmanedd	2022-12-29 16:59:54+00:00	0	Twitter for iPhone	ya sih mungkin tempat saya aga pedesaan gitu a...
4	lukmanedd	2022-12-29 16:59:53+00:00	0	Twitter for iPhone	Sebelumnya mohon maaf dengan sangat 🙏 serius n...

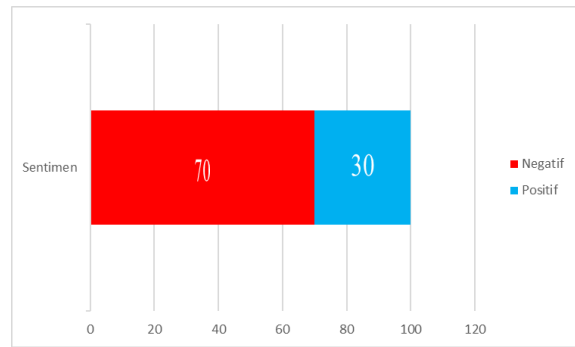
Gambar 3. Visualisasi Data

Pelabelan manual menghasilkan sentimen negatif berjumlah 70 dan sentimen positif berjumlah 30. Data ditampilkan pada Gambar 4.

User	Date Created	Number of Likes	Source of Tweet	Tweet	Sentimen	
0	el_echo	2022-12-30 06:55:33	2	Twitter for Android	wing2 lagak'e paling teranyaya pas tv analog ...	Negatif
1	ibkindartan48	2022-12-30 06:51:13	7	Twitter for Android	Gue masih menggunakan TV analog bukan TV digita...	Positif
2	delikinet	2022-12-30 06:47:28	0	Echobox	Pemadaman siaran TV analog dan dialihkan ke TV...	Negatif
3	lukmanedd	2022-12-29 23:59:54	0	Twitter for iPhone	ya sih mungkin tempat saya aga pedesaan gitu a...	Negatif
4	lukmanedd	2022-12-29 23:59:53	0	Twitter for iPhone	Sebelumnya mohon maaf dengan sangat 🙏 serius n...	Negatif

Gambar 4. Visualisasi Data Pelabelan Sentimen

Pada data, jumlah label sentimen ditampilkan dengan visual diagram batang. Dapat dilihat pada Gambar 5.



Gambar 5. Perbandingan Label Negatif dan Positif

Pada Gambar 5 dijelaskan bahwa sentimen yang berlabel Negatif lebih mendominasi terhadap sentimen yang berlabel Positif. Terkait hal itu menunjukkan bahwa pengguna Twitter tidak setuju terkait pemberhentian tv analog yang membuat sentimen berlabel Negatif lebih mendominasi. Kata yang sering muncul dapat melihat pada Gambar 6.



Gambar 6. Wordcloud Kemunculan Kata

Kemunculan kata pada data berlabel Negatif dapat dilihat pada Gambar 7.



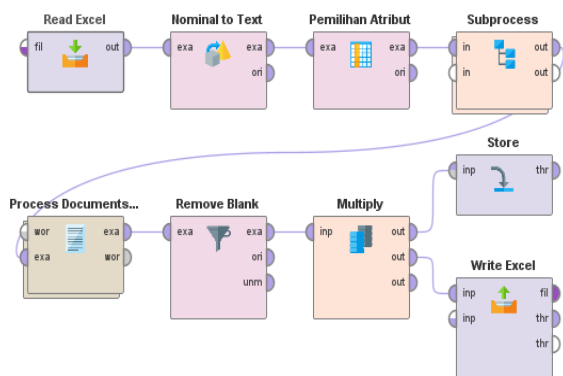
Gambar 7. Wordcloud Kemunculan Kata (Negatif)

Kemunculan kata pada data berlabel Positif dapat dilihat pada Gambar 8.



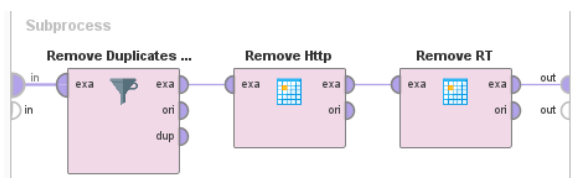
**Gambar 8.** Wordcloud Kemunculan Kata (Positif)

*Preprocessing* dilakukan pada data atribut ‘*Tweet*’ guna membersihkan data agar data menjadi bersih. Atribut yang digunakan hanya ‘*Tweet*’ dan ‘*Sentimen*’ dan data bersih akan disimpan dengan nama file ‘*databersihvanalog.xlsx*’. Proses ini menggunakan *tools* RapidMiner. *Preprocessing* pada *tools* RapidMiner dapat dilihat pada Gambar 9.



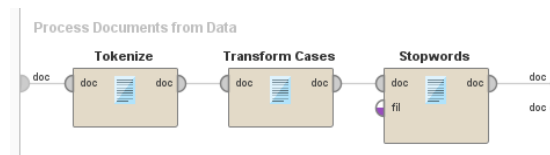
**Gambar 9.** Preprocessing Text

Pada *Subprocess* memiliki beberapa tahapan. Tahapan yang dilalui berupa *Remove Duplicates* untuk menghapus data yang ganda atau data yang sama, *Remove Http (Replace1)* untuk menghilangkan pranala yang ada, dan *Remove RT (Replace2)* untuk menghilangkan *retweet* yang ada. Tahapan dapat dilihat pada Gambar 10.



**Gambar 10.** Subprocess

Pada *process documents from data* memiliki beberapa tahapan. Tahapan yang dilalui berupa *tokenize* untuk pemisahan kalimat menjadi kata, *transform cases* untuk pengubahan huruf menjadi kecil semua, dan *filter stopwords* untuk pembuangan kata yang tidak penting. Tahapan dapat dilihat pada Gambar 11.



**Gambar 11.** Process Document From Data

Selanjutnya digunakan *Remove Blank* untuk menghilangkan data yang kosong. Data yang sudah bersih disimpan kedalam *xlsx* dan data pada RapidMiner. Selanjutnya data bersih digunakan untuk menentukan model terbaik.

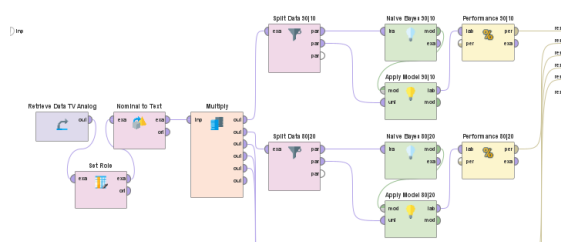
### 3.4 Modelling

Model yang digunakan yaitu *Naive Bayes* untuk pengklasifikasian terhadap sentimen pemberhentian tv analog. Data bersih digunakan untuk pembuatan model terbaik dengan melakukan lima pengujian pembagian data latih dan data uji sebagai berikut:

**Tabel 2.** Pengujian Pembuatan Model

Pengujian	Data Latih : Data Uji
1	50:50
2	60:40
3	70:30
4	80:20
5	90:10

Tahapan pengujian pembuatan model *Naive Bayes* pada RapidMiner dengan lima pengujian. Tahapan tersebut dapat dilihat pada Gambar 12.



**Gambar 12.** Tahapan Pengujian Pembuatan Model

Data bersih untuk membuat model *Naive Bayes* dengan lima pengujian yang ada pada Tabel 2. Hasil pengujian pembuatan model yang menggunakan RapidMiner dapat dilihat pada tabel berikut:

**Tabel 3.** Hasil Pengujian Model *Naive Bayes*

Pengujian	Data Latih: Data Uji	Data Latih	Data Uji	Akurasi
1	50:50	50	50	58,00%
2	60:40	60	40	67,50%
3	70:30	70	30	70,00%
4	80:20	80	20	80,00%
5	90:10	90	10	90,00%

Berdasarkan Tabel 3 model terbaik terdapat pada pengujian 5 dengan pembagian data latih dan



data uji 90:10 yang menghasilkan akurasi 90%. Selanjutnya model terbaik akan dievaluasi dengan menggunakan *confusion matrix*.

### 3.5 Evaluation

Model terbaik akan dievaluasi pada tahap ini. Model terbaik terdapat pada Pengujian 5 dengan hasil akurasi yang didapatkan sebesar 90% dengan pembagian data latih dan data uji 90:10. Model terbaik diterapkan kedalam *confusion matrix* untuk mengevaluasi model dan mendapatkan hasil akurasi, presisi, dan *recall*. Hasil *confusion matrix* dapat dilihat pada Gambar 13.

	true Negatif	true Positif	class precision
pred. Negatif	6	0	100.00%
pred. Positif	1	3	75.00%
class recall	85.71%	100.00%	

Gambar 13. Hasil *Confusion Matrix*

Hasil *confusion matrix* disajikan dengan tabel. Tabel 4 menjelaskan hasil dari *confusion matrix* pada model terbaik pengujian 5 dengan data latih berjumlah 90 data dan data uji berjumlah 10 data.

Tabel 4. Performa Algoritma *Naive Bayes*

Performa	Rumus	Perhitungan	Hasil (%)
Akurasi	$\frac{TP+TN}{(TP+TN+FP+FN)}$	$\frac{3+6}{(3+6+1+0)}$	90
Presisi	$\frac{TP}{(TP+FP)}$	$\frac{3}{(3+1)}$	75
Recall	$\frac{TP}{(TP+FN)}$	$\frac{3}{(3+0)}$	100

Hasil performa dari model terbaik mendapatkan akurasi 90%, presisi 75%, dan *recall* 100% yang menandakan model tersebut efektif. Upayakan data latih seimbang antara label satu dengan yang lainnya agar mendapatkan hasil yang terbaik dalam penerapan model *Naive Bayes*. Semakin banyak data akan mendapatkan performa yang baik pula.

## 4. KESIMPULAN

Analisis sentimen terhadap pemberhentian tv analog dengan menggunakan algoritma *Naive Bayes* dengan kata kunci 'tvanalog' dalam kurun waktu 1 Januari 2022 s.d. 30 Desember 2022 dengan batas maksimal 100 data. Didapatkan sentimen terbanyak ialah sentimen berlabel Negatif yang menandakan bahwa pengguna Twitter tidak setuju atas pemberhentian tv analog. Dilakukan lima pengujian terhadap data menggunakan algoritma *Naive Bayes* untuk mendapatkan model terbaik. Pengujian model terbaik didapatkan pada pengujian 5 dengan pembagian data latih dan data uji 90:10 dan mendapatkan hasil performa dari akurasi 90%, Presisi 75%, dan *recall* 100%. Hasil tersebut menandakan penerapan model efektif.

Saran untuk penelitian selanjutnya sebagai berikut:

1. Peneliti selanjutnya dapat menggunakan algoritma atau metode lainnya selain *Naive Bayes*, contoh: *Support Vector Machine* (SVM), *Random Forest* (RF), dan *K-Nearest Neighbor* (KNN).
2. Menambahkan data terbaru dan memberikan data latih yang cukup banyak dengan sentimen yang seimbang agar mendapatkan performa yang lebih baik.
3. Tools selanjutnya dapat menggunakan Orange dan Weka dalam proses penerapan metode yang digunakan.

## PUSTAKA

- Azimah, F., Rizky, K., Wardani, N., Komputer, F. I., & Darma, U. B. (2022). *SISTEM PENDETEKSI GEJALA AWAL COVID-19 DENGAN*. 675–688.
- Fikri, M. I., Sabrila, T. S., & Azhar, Y. (2020). Perbandingan Metode *Naive Bayes* dan *Support Vector Machine* pada Analisis Sentimen Twitter. *Smatika Jurnal*, 10(02), 71–76.  
<https://doi.org/10.32664/smatika.v10i02.455>
- Fitri, E., Yuliani, Y., Rosyida, S., & Gata, W. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma *Naive Bayes*, *Random Forest* Dan *Support Vector Machine*. *Jurnal Transformatika*, 18(1), 71.  
<https://doi.org/10.26623/transformatika.v18i1.2317>
- Halim, C., Purnomo, H. D., & Wahyono, T. (2022). *Analisis Pengelompokan Wilayah Penyebaran Covid-19 Di Indonesia Dengan Metode Clustering*. 359–372.
- Hozairi, H., Anwari, A., & Alim, S. (2021). Implementasi Orange Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Dengan Model *K-Nearest Neighbor*, *Decision Tree* Serta *Naive Bayes*. *Network Engineering Research Operation*, 6(2), 133.  
<https://doi.org/10.21107/nero.v6i2.237>
- Mufidah, Y., Noah, R., Lawalatta, L., & Bragas, N. (2022). PENGARUH TINGKAT AKURASI DALAM IDENTIFIKASI GEJALA DAN TANDA PENYAKIT PADA TANAMAN. *Suparyanto Dan Rosad (2015)*, 14.
- Nugroho, D. G., Chrisnanto, Y. H., & Wahana, A. (2016). *ANALISIS SENTIMEN PADA JASA OJEK ONLINE MENGGUNAKAN METODE NAIVE BAYES*. 156–161.
- Pintoko, B. M., & L., K. M. (2018). Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode *Naive Bayes Classifier*. *E-Proceeding of Engineering*, 5(3), 8121–8130.

- Presiden Republik Indonesia. (2020). Undang Undang Republik Indonesia Nomor 11 Tahun 2020 Tentang Cipta Kerja. *Peraturan.Bpk.Go.Id, 052692*, 1–1187.
- Ratnawati, F. (2018). Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter. *INOVTEK Polbeng - Seri Informatika*, 3(1), 50. <https://doi.org/10.35314/isi.v3i1.335>
- Widodo, S., Setiawan, D., Ridwan, T., & Ambari, R. (2022). Perancangan Deteksi Emosi Manusia berdasarkan Ekspresi Wajah Menggunakan Algoritma VGG16. *Syntax : Jurnal Informatika*, 11(01), 01–12. <https://doi.org/10.35706/syji.v11i01.6594>

HALAMAN INI SENGAJA DIKOSONGKAN