

## ANALISIS SENTIMEN MENGENAI PRODUK INOVASI INVISIBLE TV MENGUNAKAN ALGORITMA NAÏVE BAYES

### SENTIMENT ANALYSIS USING THE NAÏVE BAYES ALGORITHM FOR INVISIBLE TV INNOVATION PRODUCT

Sri Lestari<sup>1</sup>, Maryana Febryanti<sup>2</sup>

<sup>1,2</sup>Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Jakarta  
[sri.lestari1203@gmail.com](mailto:sri.lestari1203@gmail.com)<sup>1</sup>, [maryana.af18@gmail.com](mailto:maryana.af18@gmail.com)<sup>2</sup>

#### ABSTRACT

Television is one of the main sources of information for society, with increasingly sophisticated technological developments, now there is a new innovation, Invisible TV, a product that introduces the concept of television being transparent. The launch of this product caused mixed reactions from the public on social media on the YouTube platform. People give their opinions through the comment's column, this triggers pros and cons. Therefore, sentiment analysis is needed to understand and interpret people's views on Invisible TV innovation, which tends to be positive or negative. YouTube sentiment analysis is a technique used to identify someone's opinion or sentiment in comments on the platform. In this research, researchers implemented Term Frequency-Inverse Document Frequency (TF-IDF) weighting. With the final data being processed as 2186 data using the Naïve Bayes Algorithm, the Sentiment results showed that 28.41% had positive sentiment totaling 621 data while 71.59% had negative sentiment totaling 1565 data, with an accuracy result of 98.90%. Based on analysis using the RapidMiner application, the high level of negative sentiment indicates public disinterest in Invisible TV innovation.

**Keywords:** Data Mining, Sentiment Analysis, Invisible TV, Naïve Bayes.

#### ABSTRAK

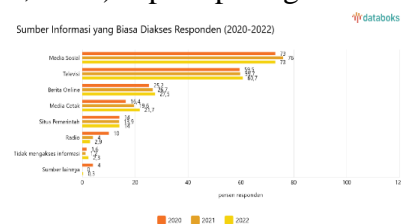
Televisi merupakan salah satu sumber informasi utama bagi Masyarakat, dengan perkembangan teknologi yang semakin canggih, kini hadir inovasi baru, *Invisible TV*, sebuah produk yang memperkenalkan konsep televisi menjadi transparan. Peluncuran produk ini menimbulkan reaksi beragam dari publik di media sosial pada platform *YouTube*. Masyarakat memberikan opini mereka melalui kolom komentar, hal ini memicu pro dan kontra. Oleh karena itu, diperlukan analisis sentimen untuk memahami dan menginterpretasikan pandangan masyarakat terhadap inovasi *Invisible TV*, lebih cenderung positif atau negatif. Analisis sentimen *YouTube* adalah teknik yang digunakan untuk mengidentifikasi pendapat atau sentimen seseorang dalam komentar di platform tersebut. Dalam penelitian ini, Peneliti mengimplementasikan pembobotan *Term Frequency-Inverse Document Frequency (TF-IDF)*. Dengan final data di proses sebanyak 2186 data menggunakan Algoritma *Naïve Bayes*, hasil Sentimen menunjukkan 28,41% bersentimen positif yang berjumlah 621 data sedangkan 71.59% bersentimen negatif yang berjumlah 1565 data, dengan hasil akurasi sebesar 98.90%. Berdasarkan analisis menggunakan aplikasi *RapidMiner*, tingginya sentimen negatif menandakan ketidaktertarikan masyarakat terhadap inovasi *Invisible TV*.

**Kata Kunci:** Data Mining, Analisis sentimen, Invisible TV, Naïve Bayes.

#### PENDAHULUAN

Televisi merupakan media elektronik yang digunakan untuk menyampaikan dan menyebarkan informasi kepada Masyarakat (Dwi Purniati et al., 2022). Pada era digital yang terus berkembang, Perusahaan teknologi terus berinovasi untuk menciptakan televisi yang semakin canggih dari tahun ke tahun yang dapat memenuhi kebutuhan dan keinginan konsumen. Televisi masih menjadi salah satu sumber informasi Masyarakat

Indonesia pada tahun 2020-2022 (Annur Mutia, 2023) seperti pada gambar berikut:



Sumber: Kementerian Komunikasi dan Informatika (Kominfo) Kantada Insight Center (KIC)

Informasi Lebih

### Gambar 1. Jumlah Persentase Sumber Informasi Yang Biasa Diakses di Indonesia (2020-2022)

Berdasarkan data yang bersumber pada <https://databoks.katadata.co.id/televi> berada di urutan ke-2 dengan jumlah persentase 59.5% tahun 2020, 59.7% tahun 2021 dan 60.7% tahun 2022. Dengan demikian memungkinkan televisi masih bisa berkembang dan digunakan di masa mendatang.

*Invisible TV* merupakan produk inovasi era 2020an sampai sekarang yang menjanjikan pengalaman menonton yang revolusioner dengan desain yang minimalis dan kemampuan untuk menyatu sempurna dengan lingkungan sekitarnya (Galeon, 2019). *Xiaomi* meluncurkan TV transparan pertama di dunia, *Mi TV LUX Transparent Edition*, pada 11 Agustus 2020 dengan harga US\$ 7.000 atau sekitar Rp 104 juta (Iskandar, 2020).

*YouTube* adalah *platform* daring yang menawarkan beragam konten visual, termasuk video klip, film dan karya-karya pengguna, tanpa dikenakan biaya (Adelia et al., 2023). Melalui fitur *comment* memberikan akses Masyarakat untuk berekspresi dan memberikan tanggapan positif ataupun negatif terhadap tayangan video *review Invisible TV brand Panasonic*.

Analisis Sentimen adalah metode yang digunakan untuk mengolah teks guna memperoleh wawasan mengenai apakah sentimen yang terkandung di dalamnya cenderung positif, netral atau negatif (Mustofa & Prasetyo, 2021). Dalam penulisan ini, menggunakan *Naïve Bayes Classifier* menggunakan konsep statistik, khususnya teori probabilitas, untuk menyelesaikan masalah dalam pembelajaran terawasi (*supervised learning*) (Mustika et al., 2021). *Naïve Bayes* dapat digunakan sebagai salah satu pendekatan untuk menganalisis sentimen terhadap *Invisible TV*.

Dalam penelitian yang dilakukan oleh Hudha, M., Supriyati, E., dan Listyorini, T., (2022) tentang Analisis Sentimen Pengguna *Youtube* Terhadap

Tayangan #Matanajwamenantiterawan Dengan Metode *Naïve Bayes Classifier* menghasilkan akurasi sebesar 90.36%. Sentimen netral lebih mendominasi dengan jumlah 1232 data netral, 90 data negatif dan 78 data positif (Hudha et al., 2022).

Menurut Parulian, M.M., Siswo, R.A.A., Rendian, S.R., (2021) tentang Analisis Sentimen Komentar Pada Saluran *Youtube Food Vlogger* Berbahasa Indonesia Menggunakan Algoritma *Naïve Bayes*, didapatkan hasil akurasi 90.60% dari data komentar sebanyak 1702 komentar yang terdiri dari 848 data positif dan 854 data negatif (Munthe et al., 2021).

Adapun penelitian yang berjudul Analisis Sentimen Penilaian Masyarakat Indonesia terhadap *GeNose* pada Komentar *Youtube* Menggunakan Metode *Naïve Bayes* oleh Jonathan, M., dan Nataliani, Y., Hasil klasifikasi menghasilkan *accuracy* mencapai 90.70% dengan kategori komentar bersifat negatif. Nilai *recall* sebesar 87.60% dan *precision* sebesar 93.39% (Jonathan & Nataliani, 2022). Dengan penjelasan di atas, Peneliti memilih judul “**Analisis Sentimen Mengenai Produk Inovasi Invisible TV Menggunakan Algoritma Naïve Bayes**”. Penelitian dengan judul tersebut menarik perhatian masyarakat serta relevan dengan tren teknologi saat ini, sehingga analisis sentimen menggunakan algoritma *Naïve Bayes* dapat memberikan wawasan dalam memahami opini konsumen dan mengidentifikasi respons positif dan negatif.

## METODE

### Data Penelitian

Data yang digunakan dalam penelitian merupakan data kualitatif. Data kualitatif adalah tipe data yang bersifat deskriptif dan umumnya tidak dapat diukur dengan angka (Subandi, 2011). Data ini lebih fokus pada karakteristik, atribut, atau kualitas dari fenomena yang diamati. Data kualitatif dibagi menjadi 2 yaitu:

- a. Data Primer

Data primer adalah data yang diperoleh langsung oleh peneliti dari sumber aslinya untuk tujuan penelitian tertentu. Data ini dikumpulkan melalui metode yang dirancang dan dijalankan oleh peneliti sendiri, sehingga hasilnya lebih sesuai dengan kebutuhan spesifik penelitian tersebut. Data Primer dalam penelitian ini ialah observasi sentimen melalui *platform YouTube*.

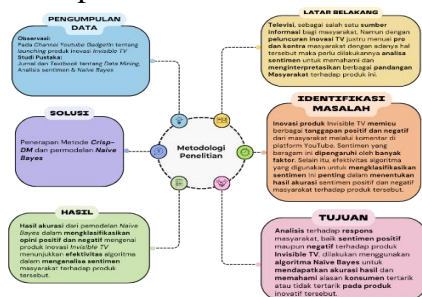
b. Data Sekunder

Data sekunder merupakan informasi yang telah dikumpulkan oleh pihak lain dan telah tersedia dalam berbagai bentuk, seperti publikasi, laporan, atau dokumen. Berikut merupakan data sekunder:

Studi pustaka melibatkan pengumpulan, membaca, analisis, dan sintesis sumber-sumber literatur yang relevan terkait dengan topik penelitian tertentu. Dalam penelitian ini Studi Pustaka berdasarkan jurnal dan *textbook* mengenai topik analisis sentimen, *data mining* dan juga *Naïve Bayes Algorithm*.

**Penerapan Metodologi**

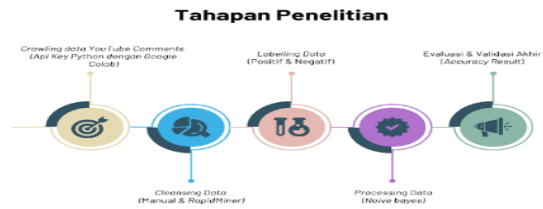
Penerapan metodologi adalah penggunaan metode dalam sebuah penelitian. Metodologi merupakan suatu rencana terstruktur yang digunakan untuk mengumpulkan data, menganalisis informasi dan mencapai tujuan tertentu dengan mengikuti langkah-langkah yang telah ditetapkan.



**Gambar 2. Tahapan Penerapan Metodologi**

**Tahapan Penelitian**

Berikut merupakan tahapan penelitian:



**Gambar 3. Tahapan Penelitian**

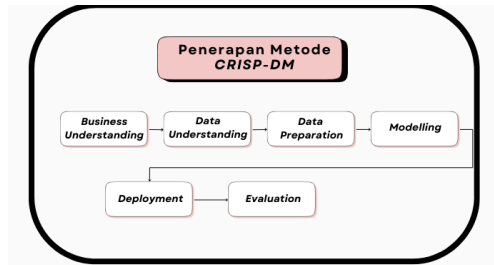
Dalam penelitian ini meliputi 4 proses yang di lakukan yaitu:

1. *Crawling data API Key YouTube*  
Objek pada deskripsi data penelitian ini adalah opini komentar *YouTube* terhadap Penayangan Video mengenai *Invisible TV*. Data komentar yang diambil berbahasa Indonesia dengan *keywords search TV* transparan.
2. *Cleansing Data*  
Komentar yang belum diolah melalui beberapa proses, yaitu: *cleaning, tokenize, transform cases, stopword* dan *filtering*.
3. *Labeling Data*  
Data yang sudah melalui proses *cleansing* kemudian melakukan pelabelan secara manual dan otomatis. Pelabelan secara manual dilakukan sebagian dan selanjutnya dilakukan oleh *RapidMiner* mengikuti pola yang sudah dibentuk maka *RapidMiner* akan merekam dan memprediksi opini mengikuti sebagian opini yang telah terisi.
4. *Processing Data*  
Pada penelitian ini akan dilakukan pengujian menggunakan pembobotan *TF-IDF* untuk melihat hasil pengujian data yang diperoleh dari tahapan *modeling* dengan menggunakan *Naïve Bayes Algorithm* untuk hasil akurasi dari algoritma tersebut.
5. *Evaluasi dan Validasi Akhir*  
Tahap terakhir penelitian ialah melihat hasil akurasi dan persentase opini posisi dan negatif.

**Rancangan Pengujian**

Pada tahapan ini pendekatan yang digunakan menggunakan metode *Cross Industry Standard for Data Mining (CRISP-DM)*. *CRISP-DM* merupakan

metode yang menggunakan model proses pengembangan data yang banyak digunakan para ahli untuk memecahkan masalah (Tuga Mauritsius, 2020). Proses penelitian ini mengacu pada enam tahap terdapat dalam *CRISP-DM* ini yakni dijelaskan sebagai berikut:



**Gambar 4. Metodologi CRISP-DM**

1. *Business Understanding* (Pemahaman Bisnis)  
Pemahaman bisnis adalah langkah awal penting dalam metodologi *CRISP-DM*. Tahap ini fokus pada pemahaman objek bisnis, metode pengumpulan data, dan penyesuaian tujuan pemodelan dengan kebutuhan bisnis, untuk memastikan model yang optimal dan relevan dengan konteks bisnis.
2. *Data Understanding* (Pemahaman Data)  
Evaluasi data bertujuan untuk mengidentifikasi potensi masalah yang bisa muncul, membentuk dasar analitik penelitian melalui ringkasan data, dan mengungkapkan masalah yang ada. Proses ini penting untuk memastikan distribusi data sesuai harapan dan untuk mengidentifikasi pola atau masalah seperti nilai yang hilang, pencilan, atau distribusi tidak merata. Masalah-masalah ini harus diatasi pada tahap Persiapan Data untuk menghindari gangguan pada tahap pemodelan selanjutnya.
3. *Data Preparation* (Persiapan Data)  
Tahap Persiapan Data fokus pada memperbaiki masalah data dan menciptakan variabel turunan agar sesuai dengan algoritma yang dipilih.
4. *Modeling*  
Tahap Pemodelan ini bertujuan untuk membuat model prediktif atau

deskriptif menggunakan teknik statistika dan Pembelajaran Mesin.

#### 5. *Evaluation*

Penafsiran hasil dari model yang telah dibuat. Evaluasi ini memastikan bahwa model sesuai dengan tujuan awal yang ditetapkan. Proses ini penting untuk memvalidasi apakah model memenuhi kebutuhan dan harapan bisnis yang telah ditentukan sebelumnya. Pada tahap ini akan dilakukan analisis atau pengukuran ketepatan terhadap pemodelan yang telah dilakukan.

#### 6. *Deployment*

Tahap Implementasi dalam *CRISP-DM* sangat penting dan direncanakan sejak tahap Pemahaman Bisnis. Implementasi mencakup cara menghasilkan nilai dari model, mengonversi skor Keputusan dan mengintegrasikan hasil ke dalam sistem operasional. Rencana ini juga mengakui bahwa model harus terus diperbarui dan disesuaikan sesuai kebutuhan.

## HASIL DAN PEMBAHASAN

### Alat Penelitian

Untuk melakukan penelitian, dibutuhkan berbagai alat yang dapat mendukung kelancaran proses penelitian tersebut. Alat-alat ini meliputi perangkat lunak (*software*) berupa *Microsoft Excel* dan *RapidMiner* serta perangkat keras (*hardware*) berupa *Lenovo Laptop-3A52L45C* dan *Printer HP Deskjet 1510 Series*.

### Implementasi Dan Pengujian

Pada penelitian ini metode yang digunakan adalah *Naïve Bayes Algorithm Classification*. Data yang digunakan ialah komentar yang di *export* menjadi data *csv* dari aplikasi media sosial *YouTube keyword search TV* Transparan pada *channel GadgetIn* menggunakan *API Key Youtube* dengan menggunakan *python* pada layanan *Google Colab* yang di ambil tanggal 16 September 2021 sampai 20 April 2024, setelah data tersebut di dapatkan lalu masuk ke proses *penerapan*

*CRISP-DM* yaitu *Business Understanding* (pemahaman bisnis), *Data Understanding* (pemahaman data), *Data Preparation* (persiapan data), *Modeling* (pemodelan), *Evaluation* (evaluasi) dan *Deployment* (pengembangan). Data tersebut akan diproses dalam pengujian algoritma *Naïve Bayes* menggunakan tools *RapidMiner* supaya dapat nilai akurasi klasifikasi yang baik dan optimal.

Pengujian ini dapat dilakukan dengan menggunakan metode *Cross-Industry Standard Process for Data Mining (CRISP-DM)* yang memiliki 6 tahapan yaitu:

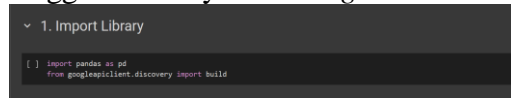
### 1. *Business Understanding* (Pemahaman Bisnis)

Televisi merupakan salah satu sumber informasi utama bagi Masyarakat, dengan perkembangan teknologi yang semakin canggih, kini hadir inovasi dengan kemampuan unik untuk menjadi transparan atau tidak terlihat ketika baru, *Invisible TV* sebuah produk yang memperkenalkan konsep televisi dengan kemampuan unik untuk menjadi transparan atau tidak terlihat ketika tidak digunakan, memberikan pengalaman visual yang futuristik dan estetis. Tujuan dasar adalah untuk mencari nilai akurasi, presisi *recall* dan menganalisa *sentiment public* tentang produk *Invisible TV*.

### 2. *Data Understanding* (Pemahaman Data)

Memahami informasi data yang akan digunakan dalam penelitian, adapun beberapa tahap untuk melakukan penelitian ini yaitu Pengumpulan data awal yaitu melakukan pengumpulan data yang dibutuhkan untuk mendukung dalam melakukan pemahaman data. Adapun sumber data utama yang digunakan dalam penelitian ini adalah data komentar TV Transparan pada channel *GadgetIn* pada tanggal 16 September 2021 sampai 20 April 2024. Sebelum melakukan proses *crawling* data perlu melakukan pendaftaran dan aktivasi *YouTube Data API*. Login ke

*Google Developer Console* dengan akun *Google* kemudian membuat *project* baru kemudian lengkapi data yang perlu diisi. Langkah selanjutnya aktifkan layanan *API* pada halaman *project* dan cari *YouTube data API V3*. Kemudian dari halaman *dashboard*, buat kredensial (*create Credential*) agar *API* tersebut dapat digunakan dan lengkapi *form* isian. Setelah itu akan mendapatkan *API Key* yang nantinya akan dimasukkan ke dalam Bahasa pemrograman *Python* untuk *crawling* data. Berikut merupakan proses *crawling* data dengan *Python* menggunakan layanan *Google Colab*.



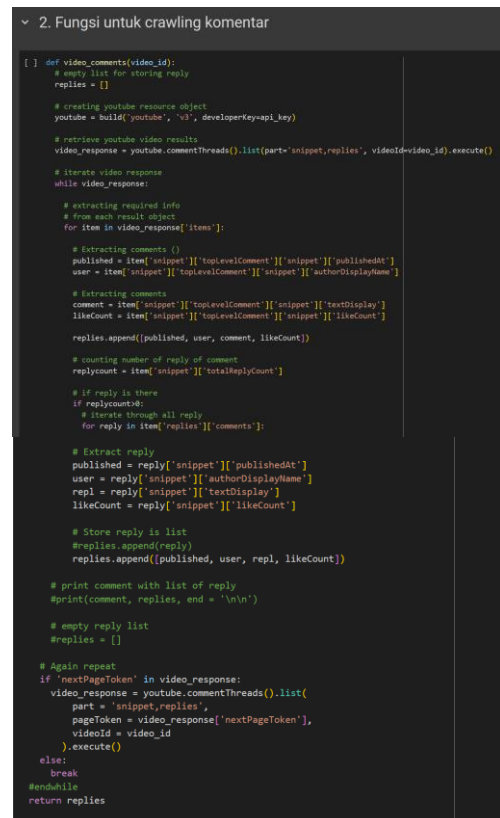
```

1. Import Library

[ ] import pandas as pd
from googleapiclient.discovery import build

```

Gambar 5. *Import Library* Proses *Crawling* Data



```

2. Fungsi untuk crawling komentar

[ ] def video_comments(video_id):
    # empty list for storing reply
    replies = []

    # creating youtube resource object
    youtube = build('youtube', 'v3', developerKey=api_key)

    # retrieve youtube video results
    video_response = youtube.commentThreads().list(part='snippet,replies', videoId=video_id).execute()

    # iterate video response
    while video_response:
        # extracting required info
        # from each result object
        for item in video_response['items']:
            # Extracting comments ()
            published = item['snippet']['topLevelComment']['snippet']['publishedAt']
            user = item['snippet']['topLevelComment']['snippet']['authorDisplayName']

            # Extracting comments
            comment = item['snippet']['topLevelComment']['snippet']['textDisplay']
            likeCount = item['snippet']['topLevelComment']['likeCount']

            replies.append((published, user, comment, likeCount))

        # counting number of reply of comment
        replyCount = item['snippet']['totalReplyCount']

        # if reply is there
        if replyCount > 0:
            # iterate through all reply
            for reply in item['replies']['comments']:
                # Extract reply
                published = reply['snippet']['publishedAt']
                user = reply['snippet']['authorDisplayName']
                repl = reply['snippet']['textDisplay']
                likeCount = reply['snippet']['likeCount']

            # Store reply in list
            #replies.append(reply)
            replies.append((published, user, repl, likeCount))

        # print comment with list of reply
        #print(comment, replies, end = '\n\n')

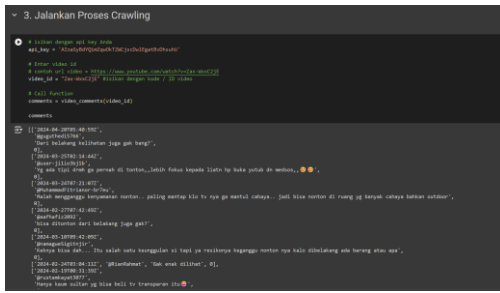
        # empty reply list
        #replies = []

    # Again repeat
    if 'nextPageToken' in video_response:
        video_response = youtube.commentThreads().list(
            part = 'snippet,replies',
            pageToken = video_response['nextPageToken'],
            videoId = video_id
        ).execute()
    else:
        break
    #return replies
    return replies

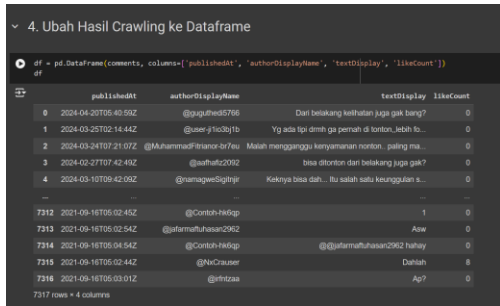
```

Gambar 6. Fungsi Untuk Proses *Crawling* Data





Gambar 7. Menjalankan Fungsi Untuk Proses Crawling Data



Gambar 8. Ubah Hasil Crawling Data Ke Data frame



Gambar 9. Hasil Tersimpan Ke File CSV



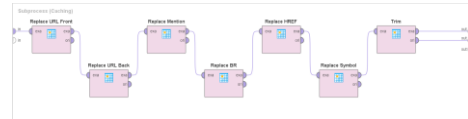
Gambar 10. Hasil Pengumpulan Data

3. Data Preparation (Persiapan data)

a. Cleansing data

Pada tahapan ini peneliti akan melakukan proses *Cleansing* dengan menggunakan *RapidMiner* sebelum melakukan pemberian label pada sentimen, agar *dataset* pada *file csv* terbaru terhindar dari duplikasi data dan tanda yang tidak diperlukan. Pada proses ini akan dilakukan pembersihan dari berbagai *noise* seperti menghilangkan *link URL*, *mention*, *symbol*, karakter dan menghilangkan spasi awal, akhir dari nilai nominal dan lainnya. Setelah selesai menjalankan proses *Cleansing* dengan menggunakan *RapidMiner*. Awalnya terdapat 7.318

data pada *file csv*, kemudian setelah melewati proses *Cleansing* maka data komentar berkurang menjadi 7.002 data. Selanjutnya Dilakukan pengecekan manual, data yang didapatkan masih terdapat komentar yang *noise*, komentar yang tidak pantas dan tidak sesuai dengan topik maka dilakukan *cleansing* manual sehingga data yang berhasil didapatkan sebanyak 2260 data.



Gambar 11. Proses Cleansing Data

Gambar 12. Hasil Cleansing Data RapidMiner

Gambar 13. Hasil Cleansing Data Manual

b. Labeling data

Tahap berikutnya adalah melakukan klasifikasi atau menentukan sentimen dari setiap komentar yang ada. Dalam proses ini, peneliti memberikan label sentimen secara manual pada 678 komentar yang akan digunakan sebagai data latih. Setelah itu, sebanyak 1582 komentar akan diklasifikasikan secara otomatis berdasarkan sentimen menggunakan perangkat lunak *RapidMiner*.



Gambar 14. labeling data Manual

c. Pre-Processing data

Tahap *Preprocessing* data bertujuan untuk mengubah data mentah menjadi data yang siap diolah, adapula tahap ini dibagi menjadi beberapa bagian, yaitu:

1. *Tokenizing*, yaitu tahapan untuk melakukan pemisahan kata dalam suatu kalimat dengan tujuan untuk proses analisis teks lebih lanjut, seperti teks “mengganggu kenyamanan menonton” setelah melewati tahap *Tokenizing*, akan menjadi 4 kata, yaitu “mengganggu”, “kenyamanan” dan “menonton”.
2. *Transform Cases*, yaitu tahapan untuk mengubah data komentar menjadi *lower case* (huruf kecil), seperti teks “Bagus” menjadi “bagus”.
3. *Stemming*, yaitu tahapan untuk proses mereduksi kata berimbuhan menjadi kata dasar, menghilangkan imbuhan awalan dan akhiran yang terdapat dalam teks seperti “membeli” menjadi “beli”.
4. *Filter Tokens (by Length)*, yaitu tahapan untuk mengeliminasi kata-kata yang kurang dari 2 huruf dan kata-kata yang melebihi 25 huruf.
5. *Filter Stopwords*, yaitu tahapan kata yang diabaikan dalam pemrosesan dan biasanya disimpan di dalam *stop lists*. *Stop list* ini berisi daftar kata umum yang mempunyai fungsi tapi tidak mempunyai arti, seperti “saya”, “dia”, “aku” dan “mereka”.

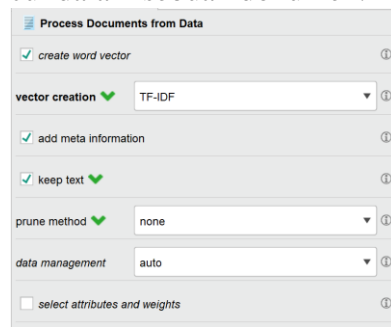


Gambar 15. Pre-Processing

d. Pembobotan Kata

Pada tahap ini, hasil *pre-processing* akan diproses lebih lanjut agar setiap

kata mendapatkan bobot tertentu. Penulis menggunakan algoritma *TF-IDF* untuk *TF-IDF* (*Pembobotan Kata*), yang merupakan singkatan dari *Term Frequency-Inverse Document Frequency*, adalah metode algoritma yang berguna untuk menghitung bobot setiap kata yang sering digunakan. Metode ini dikenal efisien, mudah diterapkan dan memberikan hasil yang akurat. Algoritma ini menghitung nilai *Term Frequency (TF)* dan *Inverse Document Frequency (IDF)* untuk setiap token (kata) di setiap dokumen dalam korpus. Secara sederhana, metode *TF-IDF* digunakan untuk mengetahui seberapa sering suatu kata muncul dalam sebuah dokumen.



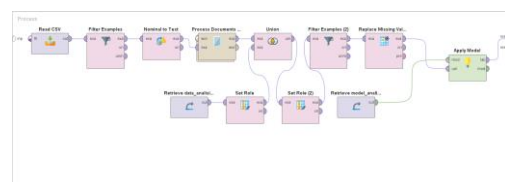
Gambar 16. Pembobotan Kata

Row No.	Sentimen	prediction...	confidence...	confidence...	text	kejar	sewa	contoh	seor
1	Negatif	Negatif	1	0	sewa	0	0	0	0
2	Positif	Positif	0	1	sewa	0	0	0	0
3	Negatif	Negatif	1	0	sewa	0	0	0	0
4	Negatif	Negatif	1	0	sewa	0	0	0	0
5	Negatif	Negatif	1	0	sewa	0	0	0	0
6	Negatif	Negatif	1	0	sewa	0	0	0	0
7	Negatif	Negatif	1	0	sewa	0	0	0	0
8	Negatif	Negatif	1	0	sewa	0	0	0	0
9	Negatif	Negatif	1	0	sewa	0	0	0	0
10	Negatif	Negatif	1	0	sewa	0	0	0	0
11	Negatif	Negatif	1	0	sewa	0	0	0	0
12	Negatif	Negatif	1	0	sewa	0	0	0	0
13	Negatif	Negatif	1	0	sewa	0	0	0	0
14	Negatif	Negatif	1	0	sewa	0	0	0	0
15	Negatif	Negatif	1	0	sewa	0	0	0	0
16	Negatif	Negatif	1	0	sewa	0	0	0	0
17	Negatif	Positif	0	1	sewa	0	0	0	0
18	Negatif	Negatif	1	0	sewa	0	0	0	0

Gambar 17. Hasil Pembobotan Kata



Gambar 18. Proses Data Latih



Gambar 19. Proses Data Uji

Row No.	text	prediction(sentiment)
1	laman tourism	Negatif
2	hadat beli tv nonton berita metro tv	Negatif
3	hadat ya bagus	Positif
4	reneh bagus	Positif
5	masak tv transparan mendung tv food ngepas obsekyang tv	Negatif
6	hadat ngepas ngepas gajay	Positif
7	hadat ngepas ngepas gajay	Positif
8	urifadon ngepas best	Positif
9	hp	Positif
10	hanga ralu ya mendung beli rumah paku ya masuk akai	Negatif
11	idaman	Positif
12	imgidat pas dlujudkan	Positif
13	implementasi tv obayangkan cocok	Positif
14	lirous akai	Negatif
15	akum mendung mendung	Negatif
16	hapa beresah gampang dlat	Negatif
17	masa ya ngepas ngepas ngepas	Positif
18	bagayay tempat	Positif

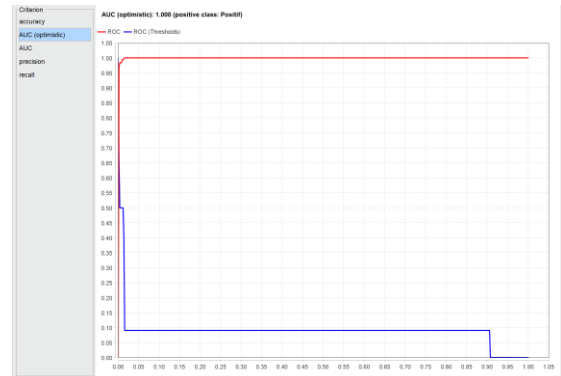
Gambar 20. Hasil Labeling Data RapidMiner

Criterion	Value
accuracy	98.90%
AUC (optimistic)	1.000 (positive class: Positif)
precision	99.34% (positive class: Positif)
recall	96.78% (positive class: Positif)

	true Negatif	true Positif	class precision
pred. Negatif	1561	20	98.73%
pred. Positif	4	601	99.34%
class recall	99.74%	96.78%	

Gambar 22. Hasil Proses Akurasi



Gambar 23. Grafik Area Under Curve (AUC)

#### 4. Modeling (Pemodelan)

Pada tahapan sebelumnya data latih dan data uji sudah melakukan *pre-processing* dan (pembobotan kata) *TF-IDF* sehingga pada tahapan ini data yang didapatkan sudah bersih dan bebas dari *noise*, jumlah data komentar didapatkan menjadi 2.186 data.



Gambar 21. Proses Modeling

Criterion	Value
accuracy	98.90%
AUC (optimistic)	1.000 (positive class: Positif)
precision	99.34% (positive class: Positif)
recall	96.78% (positive class: Positif)

	true Negatif	true Positif	class precision
pred. Negatif	1561	20	98.73%
pred. Positif	4	601	99.34%
class recall	99.74%	96.78%	

Gambar 24. Hasil Precision

#### 5. Evaluation (Evaluasi)

Metode evaluasi pada penelitian ini nantinya akan menunjukkan metrik akurasi, presisi, *AUC* dan *recall* pada tahap pengujian.

#### 6. Deployment

Hasil penelitian setelah memasuki tahap hasil akhir pengujian dipublikasikan menjadi jurnal ilmiah dan laporan akhir skripsi.

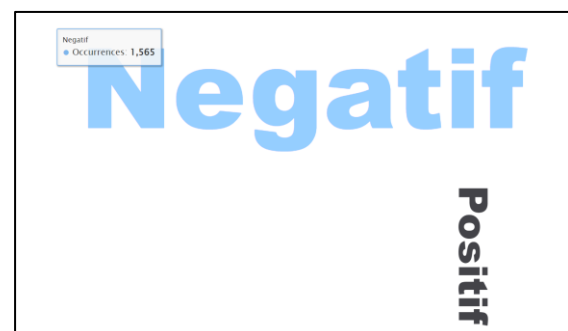
#### Hasil Akhir Pengujian

Pada penelitian ini pengoptimalan pengujian dengan *confusion matrix* untuk melihat hasil pengujian data yang diperoleh dari tahapan *modelling* dengan menggunakan *Naïve Bayes Algorithm*. Total dataset yang dikumpulkan adalah 2.186 data. Berikut ini adalah *confusion matrix* hasil dari tahapan *modelling* yang dapat dilihat hasil dari perhitungan hasil *RapidMiner* dapat dilihat pada Gambar 22. untuk menunjukkan dan membuktikan hasil klasifikasi dari model *Naïve Bayes*.

```

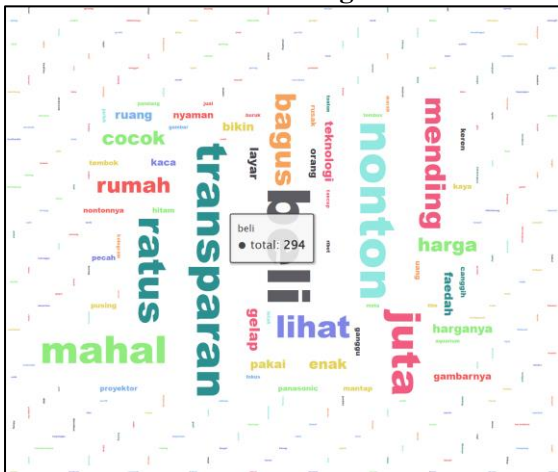
PerformanceVector
PerformanceVector:
accuracy: 98.90%
ConfusionMatrix:
True:  Negatif Positif
Negatif:  1561  20
Positif:   4    601
AUC (optimistic): 1.000 (positive class: Positif)
AUC: 1.000 (positive class: Positif)
precision: 99.34% (positive class: Positif)
ConfusionMatrix:
True:  Negatif Positif
Negatif:  1561  20
Positif:   4    601
recall: 96.78% (positive class: Positif)
ConfusionMatrix:
True:  Negatif Positif
Negatif:  1561  20
Positif:   4    601
    
```

Gambar 25. Hasil Pengujian Model Naïve Bayes





**Gambar 26. Visualisasi Hasil Wordcloud Positif dan Negatif**



**Gambar 27. Visualisasi Hasil Wordcloud per kata**

Berdasarkan pada Gambar 22. dapat dibuat kesimpulan bahwa *precision* menunjukkan tingkat ketepatan data yang diprediksi negatif terhadap banyaknya data yang benar diprediksi negatif yang menghasilkan presentase ketepatannya adalah 98.73%, sedangkan untuk data bersentimen positif memiliki *precision* sebesar 99.34%. Untuk sentimen negatif memiliki *recall* adalah 99.74% sehingga dapat disimpulkan bahwa model dapat menemukan kembali informasi atau data yang benar-benar negatif dengan sangat baik, sedangkan pada sentimen positif memiliki *recall* sebesar 96.78%. Nilai *accuracy* yang dihasilkan menggunakan model *Naïve Bayes* adalah 98.90%, dengan nilai *AUC* sebesar 1.00 sehingga dapat disimpulkan bahwa algoritma *Naïve Bayes* dapat mengklasifikasi sentimen dengan baik menggunakan data komentar TV Transparan.

## **PENUTUP**

### **Kesimpulan**

Berdasarkan penelitian yang telah dilakukan mengenai klasifikasi opini masyarakat Indonesia terhadap *Invisible TV* dapat diambil kesimpulan yaitu hasil Sentimen ditemukan 28,41% bersentimen positif yang berjumlah 621 data dan 71.59% bersentimen negatif yang berjumlah 1565 data. Maka dari hasil tersebut dapat disimpulkan bahwa

masyarakat Indonesia sebagian besar memiliki respon negatif terhadap produk *Invisible TV* dan berdasarkan klasifikasi model algoritma *Naïve Bayes*, terhadap dataset produk *Invisible TV*, didapatkan nilai *accuracy* sebesar 98.90%. Sehingga dapat dikatakan bahwa algoritma *Naïve Bayes* dapat mengklasifikasikan data secara baik.

## **Saran**

Adapun saran-saran yang disampaikan berdasarkan hasil pengamatan dan analisa selama melakukan penelitian ini adalah:

1. Diharapkan dapat dilakukan penelitian dengan topik serupa namun dengan algoritma yang berbeda atau dengan dikombinasi agar dapat dibandingkan hasil yang lebih baik.
2. Saat pembagian data latih dan data uji, dapat dilakukan dengan lebih banyak jumlah data ujinya agar memperkaya kata-kata dalam menentukan *sentiment*

## **DAFTAR PUSTAKA**

- Adelia, S., Milanda, E., Santari, J., Kesuma, D. T., Silvia, E., & Kurniawan, F. (2023). Analisis Sentimen Belajar Programming Pada Media Sosial Youtube Menggunakan Algoritma Klasifikasi Naive Bayes. *Journal of Information Technology Ampera*, 4(3), 2774–2121. <https://journal-computing.org/index.php/journal-ita/index>
- Annur Mutia, C. (2023). *Media Sosial, Sumber Informasi Utama Masyarakat Indonesia*. Databoks. <https://databoks.katadata.co.id/datapublish/2023/05/30/media-sosial-sumber-informasi-utama-masyarakat-indonesia>
- Dwi Purniati, Samsu, & Adi Iqbal. (2022). Pemanfaatan Media Televisi dalam Memenuhi Kebutuhan Informasi Masyarakat Desa Sungai Ruan Ilir Kabupaten Batang Hari. *Journal of Religion and Film*, 1(1), 19–36.

- <https://doi.org/10.30631/jrf.v1i1.3>  
 Galeon, D. (2019). "TV Tak Terlihat" Panasonic yang Baru dan Lebih Baik Menjadi Transparan Saat Dimatikan. *Futurism*. [https://futurism-com.translate.googleusercontent.com/translate/g/panasonics-new-and-improved-invisible-tv-becomes-transparent-when-its-turned-off?\\_x\\_tr\\_sl=en&\\_x\\_tr\\_tl=id&\\_x\\_tr\\_hl=id&\\_x\\_tr\\_pto=tc&\\_x\\_tr\\_hist=true](https://futurism-com.translate.googleusercontent.com/translate/g/panasonics-new-and-improved-invisible-tv-becomes-transparent-when-its-turned-off?_x_tr_sl=en&_x_tr_tl=id&_x_tr_hl=id&_x_tr_pto=tc&_x_tr_hist=true)
- Hudha, M., Supriyati, E., & Listyorini, T. (2022). Analisis Sentimen Pengguna Youtube Terhadap Tayangan #Matanajwamentiterawan Dengan Metode Naïve Bayes Classifier. *JIKO (Jurnal Informatika Dan Komputer)*, 5(1), 1–6. <https://doi.org/10.33387/jiko.v5i1.3376>
- Iskandar. (2020). *Xiaomi Pamer TV OLED Tembus Pandang, Seperti Apa Bentuknya?* Liputan6.Com. <https://www.liputan6.com/tekno/read/4329049/xiaomi-pamer-tv-oled-tembus-pandang-seperti-apa-bentuknya?page=3>
- Jonathan, M., & Nataliani, Y. (2022). Analisis Sentimen Penilaian Masyarakat Indonesia terhadap GeNose pada Komentar Youtube Menggunakan Metode Naïve Bayes. *Jurnal Matematika Dan Aplikasi*, 11(01). <https://ejournal.unsrat.ac.id/index.php/decartesian/article/view/38339>
- Munthe, M. P., Ansori, A. S. R., & ... (2021). Analisis Sentimen Komentar Pada Saluran Youtube Food Vlogger Berbahasa Indonesia Menggunakan Algoritma Naïve Bayes. *EProceedings of Engineering*, 8(6), 11909–11916. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/16897>
- Mustika, Ardila, Y., Manuhutu, A., Ahmad, N., Hasbi, I., Guntoro, Manuhutu, M. A., Ridwan, M., Hozairi, Wardhani, A. K., Alim, S., Romli, I., Religia, Y., Octafian, D. T., Sufandi, U. U., & Ernawati, I. (2021). Data Mining dan Aplikasinya. In *Penerbit Widina*. <https://repository.penerbitwidina.com/uk/publications/351768/data-mining-dan-aplikasinya>
- Mustofa, R. L., & Prasetyo, B. (2021). Sentiment analysis using lexicon-based method with naive bayes classifier algorithm on #newnormal hashtag in twitter. *Journal of Physics: Conference Series*, 1918(4). <https://doi.org/10.1088/1742-6596/1918/4/042155>
- Subandi. (2011). Qualitative Description as one Method in Performing Arts Study. *Harmonia*, 19, 173–179.
- Tuga Mauritsius, F. B. (2020). *Cross-Industry Standard Process for Data Mining (CRISP-DM)*. <https://mmsi.binus.ac.id/2020/09/18/cross-industry-standard-process-for-data-mining-crisp-dm/>