



Tersedia Online : <http://e-journals.unmul.ac.id/>

ADOPSI TEKNOLOGI DAN SISTEM INFORMASI (ATASI)

Alamat Jurnal : <http://e-journals2.unmul.ac.id/index.php/atasi/index>



Analisis Sentimen Pengguna Youtube Terhadap Uang Baru Tahun Emisi 2022 Menggunakan Metode Naïve Bayes Classifier

Aji Akbar Mirinda Putra ^{1)*}, Islamiyah ²⁾, Muhammad Labib Jundillah ³⁾

Program Studi Sistem Informasi, Fakultas Teknik, Universitas Mulawarman

E-Mail : aj14kbar@gmail.com ¹⁾; islamiyahunmul@gmail.com ²⁾; muhammadjundillah@ft.unmul.ac.id ³⁾;

ARTICLE INFO

Article history:

Received : 28 June 2024

Revised : 29 June 2024

Accepted : 20 Februari 2024

Available online : 30 June 2024

Keywords:

Rupiah,

YouTube,

Sentiment Analysis,

Naïve Bayes Classifier,

New Currency For The 2022

Emission Year

ABSTRACT

Money is a vital commodity in economic activities and in 2022 Bank Indonesia launched the 2022 Emission Year Rupiah with a more attractive design and improved security features. In this case social media, particularly YouTube became an important platform for Indonesians to express their opinions on the change. The Naïve Bayes Classifier method is used to classify user sentiment into positive or negative categories. This research aims to analyze YouTube users' sentiment towards the 2022 Rupiah Emission Year by using the Naïve Bayes Classifier method. The main objective is to identify whether user sentiment is positive or negative towards the 2022 Rupiah Emission Year and measure the accuracy, precision, recall, and f1-score of the sentiment analysis. The analysis shows that 55.1% of YouTube users responded positively to the 2022 Rupiah Emission Year, while 44.9% responded negatively. The performance evaluation of the model resulted in 80% accuracy, 87% precision, 73% recall, and 80% f1-score. These results show that the model performs well in classifying user sentiment.

ABSTRAK

Uang merupakan komoditas vital dalam kegiatan ekonomi dan pada tahun 2022 Bank Indonesia meluncurkan Uang Rupiah Tahun Emisi 2022 dengan desain yang lebih menarik dan fitur keamanan yang lebih baik. Dalam hal ini media sosial, khususnya YouTube menjadi platform penting bagi masyarakat Indonesia untuk mengekspresikan pendapat mereka tentang perubahan tersebut. Metode Naïve Bayes Classifier digunakan untuk mengklasifikasikan sentimen pengguna ke dalam kategori positif atau negatif. Penelitian ini bertujuan untuk menganalisis sentimen pengguna YouTube terhadap Tahun Emisi Rupiah 2022 dengan menggunakan metode Naïve Bayes Classifier. Tujuan utamanya adalah untuk mengidentifikasi apakah sentimen pengguna positif atau negatif terhadap Tahun Emisi Rupiah 2022 dan mengukur *accuracy*, *precision*, *recall*, dan *f1-score* dari analisis sentimen tersebut. Hasil analisis menunjukkan bahwa 55,1% pengguna YouTube memberikan respon positif terhadap Uang Rupiah Tahun Emisi 2022, sementara 44,9% memberikan respon negatif. Evaluasi kinerja model menghasilkan *accuracy* 80%, *precision* 87%, *recall* 73%, dan *f1-score* 80%. Hasil ini menunjukkan bahwa model ini memiliki kinerja yang baik dalam mengklasifikasikan sentimen pengguna.

Kata Kunci:

Rupiah,

YouTube,

Analisis Sentimen,

Naïve Bayes Classifier,

Uang Baru Tahun Emisi 2022

APA style in citing this article:

Aji Akbar Mirinda Putra, Islamiyah & Muhammad Labib Jundillah.

Analisis Sentimen Pengguna

Youtube Terhadap Uang Baru

Tahun Emisi 2022 Menggunakan

Metode Naïve Bayes Classifier .

Adopsi Teknologi Dan Sistem

Informasi (ATASI), 3(1). Retrieve

from

<https://doi.org/10.30872/atasi.v3i1>

177

2024 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.

*) Corresponding Author

<https://doi.org/10.30872/atasi.v3i1.1177>

2024 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.

1. PENDAHULUAN

Uang berfungsi sebagai alat pertukaran dan ukuran kekayaan karena sangat penting bagi dinamika ekonomi. Memudahkan transaksi jual beli adalah fungsi utama uang. Semua orang di dunia menggunakan uang, baik dalam bentuk koin maupun kertas, termasuk mata uang Indonesia, Rupiah. Bank Indonesia menanggapi evolusi waktu dengan meluncurkan Rupiah Kertas untuk Tahun Emisi 2022 yang meningkatkan mutu dan reliabilitas mata uang tersebut. Perayaan HUT ke-77 Kemerdekaan Indonesia adalah kesempatan penting untuk memperkenalkan desain yang lebih menarik, fitur keselamatan yang lebih canggih, dan peningkatan ketahanan material dengan tujuan meningkatkan kebanggaan sebagai tanda kedaulatan Negara Kesatuan Republik Indonesia (NKRI) (Bank Indonesia, 2022).

Perkembangan teknologi juga mempengaruhi cara orang berinteraksi dan berkomunikasi satu sama lain. Media sosial, terutama YouTube muncul sebagai platform penting untuk menyampaikan pendapat dan emosi. Masyarakat Indonesia secara aktif menggunakan berbagai platform media sosial, termasuk YouTube, untuk menyampaikan pendapat mereka tentang Uang Baru Tahun Emisi 2022. Sebagai platform berbagi video *online*, YouTube menyediakan berbagai pandangan yang dapat diakses oleh siapa pun yang memiliki koneksi *internet*. Komentar yang ada di platform ini dapat digunakan sebagai sumber data penting untuk mempelajari bagaimana masyarakat melihat perubahan tersebut (Tutiasri et al., 2020).

Dalam hal ini, metode sentimen *Naïve Bayes Classifier* menjadi alat yang efektif. Meskipun sederhana, metode ini menunjukkan tingkat akurasi dan kinerja yang tinggi dalam mengklasifikasikan teks. Sebagai pendekatan probabilistik, *Naïve Bayes Classifier* digunakan untuk memilah opini pengguna YouTube terkait Uang Baru Tahun Emisi 2022 ke dalam kategori positif atau negatif (Suryani et al., 2019).

2. TINJAUAN PUSTAKA

A. Analisis Sentimen

Analisis sentimen adalah bidang yang berhubungan dengan penilaian, tanggapan, dan perasaan yang berasal dari teks yang banyak digunakan di bidang-bidang seperti *data mining*, *web mining*, dan *social media analytics* karena sentimen adalah karakteristik yang paling penting untuk menilai perilaku manusia (Chakraborty et al., 2019).

B. YouTube

YouTube sebagai situs web dan aplikasi yang sangat populer dengan menggunakan *internet* sebagai media untuk menampilkan kontennya. Pengguna memiliki kemampuan untuk mengunggah video mereka sendiri, berinteraksi dengan video tersebut, dan memberikan komentar. Meskipun YouTube telah berkembang menjadi *database* video paling populer di dunia, popularitasnya juga berkontribusi pada potensi penyalahgunaan (Hendika Permana, 2021). YouTube API merupakan layanan yang disediakan untuk pengembang, YouTube API memainkan peran penting dalam memfasilitasi integrasi dengan sumber daya video di platform ini dan memberikan pengembang kendali atas interaksi pengguna dan video. Mengintegrasikan API dengan platform ini memungkinkan pengembang membuat aplikasi dan layanan yang lebih kompleks dan meningkatkan pengalaman pengguna melalui pemanfaatan suasana yang lebih baik (Yosef et al., 2020).

C. Text Preprocessing

Text preprocessing melibatkan *case folding*, *cleaning*, *tokenizing*, dan *stemming* untuk memproses data agar dapat diklasifikasikan dengan baik (Mas Pintoko & Muslim, 2018).

D. Term Frequency–Inverse Document Frequency (TF-IDF)

Term Frequency–Inverse Document Frequency (TF-IDF) adalah metode untuk memberi bobot kata terhadap dokumen. Metode ini menghitung bobot berdasarkan seberapa sering kata itu muncul di dokumen dan seberapa sering kata itu muncul dalam kumpulan dokumen. Jika kata itu muncul sering di dokumen, maka bobotnya akan tinggi (Ramadhan et al., 2023). Perhitungan TF-IDF dapat ditunjukkan pada persamaan (1).

$$W_{dt} = TF_{dt} \times \log \left(\frac{d}{df} \right) \dots \dots \dots (1)$$

Keterangan:

W_{dt} = Bobot dokumen ke-d terhadap kata ke-t

TF_{dt} = Banyaknya kata yang dicari pada sebuah dokumen

IDF_{ft} = *Inversed Document Frequency* $\left(\frac{N}{df} \right)$

N = Total dokumen

df = Banyak dokumen yang mengandung kata yang dicari

E. Naïve Bayes Classifier

Naïve Bayes Classifier adalah metode klasifikasi yang berdasarkan *teorema bayes* dan independensi prediktor. Metode ini menghitung probabilitas dari data historis dan mengabaikan keterkaitan fitur dalam kelas. Metode ini efisien karena hanya memerlukan data pelatihan dan varian variabel dalam kelas untuk klasifikasi (Zainal Macfud et al., 2023). Rumus *Naïve Bayes Classifier* dalam secara umum dapat ditunjukkan pada persamaan (2).

$$P(H|X) = (P(H) \frac{P(H)}{P(X)}) \dots\dots\dots(2)$$

Keterangan:

- X = Data dengan *class* yang belum diketahui.
- H = Hipotesis data X merupakan suatu *class* spesifik.
- P(H|X) = Probabilitas hipotesis H berdasarkan kondisi x (*posteriori probability*)
- P(H) = Probabilitas hipotesis H (*prior probability*)
- P(X|H) = Probabilitas X berdasarkan kondisi tersebut
- P(X) = Probabilitas dari X

F. Confusion Matrix

Confusion matrix adalah matriks yang menunjukkan kemampuan sistem klasifikasi dalam *supervised learning*. Matriks ini memiliki kolom untuk kelas prediksi dan baris untuk kelas aktual. *Confusion matrix* terdiri dari empat istilah: TP (*True Positive*), TN (*True Negative*), FP (*False Negative*), dan FN (*False Positive*). Istilah-istilah ini menggambarkan kesesuaian atau ketidaksesuaian antara nilai prediksi dan nilai aktual (Saiyar, 2018). Berikut metrik evaluasi untuk *confusion matrix*:

a. *Accuracy*

Accuracy adalah persentase kasus yang diklasifikasikan dengan benar oleh model (Xu et al., 2020). Nilai *accuracy* dapat diperoleh dengan persamaan (3).

$$Accuracy = \frac{TN+TP}{TN+FP+FN+TP} \dots\dots\dots(3)$$

Keterangan:

- TP = *True Positive*
- TN = *True Negative*
- FP = *False Positive*
- FN = *False Negative*

b. *Precision*

Precision adalah rasio antara kasus positif yang diprediksi dengan benar dan semua kasus yang diprediksi positif oleh model (Xu et al., 2020). Nilai *precision* dapat diperoleh dengan persamaan (4).

$$Precision = \frac{TP}{TP+FP} \dots\dots\dots(4)$$

Keterangan:

- TP = *True Positive*
- FP = *False Positive*

c. *Recall*

Recall adalah rasio antara kasus positif yang diprediksi dengan benar dan semua kasus positif yang sebenarnya (Xu et al., 2020). Nilai *recall* dapat diperoleh dengan persamaan (5).

$$Recall = \frac{TP}{TP+FN} \dots\dots\dots(5)$$

Keterangan:

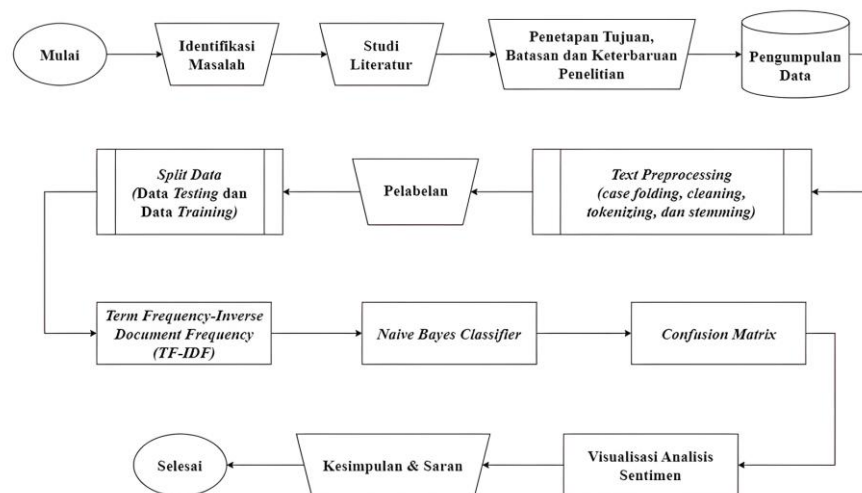
- TP = *True Positive*
- FN = *False Negative*

d. *F1-score*

F1-score adalah ukuran yang menggabungkan *precision* dan *recall* dengan cara menghitung rata-rata harmonik dari keduanya (Xu et al., 2020). Nilai *f1-score* dapat diperoleh dengan persamaan (6).

$$F1-score = 2 \times \frac{precision \times recall}{precision + recall} \dots\dots\dots(6)$$

3. METODE PENELITIAN



Gambar 1. Alur Metode Penelitian

G. Pengumpulan Data

Sumber data komentar YouTube dalam penelitian ini adalah dua video, yaitu video dari *channel* Bank Indonesia yang berjudul “Uang Rupiah Kertas Emisi 2022” dan video dari *channel* KOMPASTV yang berjudul “Prosesi Peluncuran Tujuh Uang Rupiah Kertas Tahun Emisi 2022, dari Rp 1.000 hingga Rp 100.000”. Data diambil pada tanggal 1 Mei 2023 menggunakan YouTube API dengan memasukkan *video_id* dari kedua video tersebut. Setelah proses pengambilan data selesai, data dari kedua video tersebut disatukan, lalu data komentar duplikat dihapus, dan kemudian disimpan dalam sebuah *file* dengan format CSV.

H. Text Preprocessing

Komentar-komentar YouTube yang telah terkumpul akan diproses dengan beberapa tahap pengolahan teks, yang meliputi proses *case folding*, *cleaning*, *tokenizing*, dan *stemming*. *Case folding* adalah proses mengubah semua huruf menjadi huruf kecil untuk menghindari perbedaan penulisan. *Cleaning* adalah proses menghapus karakter yang tidak relevan, seperti tanda baca, angka, emoji, dan spasi berlebih. *Stemming* adalah proses mengubah kata-kata menjadi bentuk dasarnya dengan menghilangkan imbuhan. *Tokenizing* adalah proses memecah teks menjadi kata-kata yang disebut token.

I. Pelabelan

Proses penentuan sentimen pada komentar-komentar YouTube akan dilakukan dengan memberikan label positif atau negatif pada setiap kalimat secara manual oleh seorang pakar bahasa, yaitu seorang guru bahasa Indonesia SMA Islam Bunga Bangsa Samarinda. Sentimen positif merujuk pada teks yang mengandung makna yang baik, segar, atau memuaskan. Sentimen negatif, merujuk pada teks yang mengandung makna yang buruk atau mengganggu. Data akan diberi label numerik, yaitu 1 untuk sentimen positif dan 0 untuk sentimen negatif.

J. Split Data

Pemisahan data dilakukan untuk membagi *dataset* menjadi dua bagian, yaitu data latih (*training*) dan data uji (*testing*). Data latih berperan untuk melatih model klasifikasi, sementara data uji digunakan untuk menguji dan mengevaluasi kinerja model klasifikasi yang telah dilatih pada tahap data latih. Pemisahan data ini akan dilakukan dengan proporsi 70% untuk data latih dan 30% untuk data uji.

K. Term Frequency–Inverse Document Frequency (TF-IDF)

Term Frequency–Inverse Document Frequency (TF-IDF) digunakan untuk memberikan bobot di mana setiap baris merepresentasikan satu kalimat (komentar), sementara setiap kolom merepresentasikan kata dalam seluruh kalimat. Perhitungan ini bertujuan untuk menilai seberapa relevan suatu kata dalam suatu dokumen.

L. Naïve Bayes Classifier

Pada tahap ini, dilakukan proses klasifikasi menggunakan metode *Naïve Bayes Classifier* untuk mengidentifikasi data yang termasuk dalam kategori opini positif dan negatif. Kelebihan dari metode *Naïve Bayes Classifier* terletak pada penggunaan jumlah data latih yang relatif sedikit, memungkinkan perhitungan yang lebih

simpel, cepat, dan efisien. *Naïve Bayes Classifier* digunakan pada data latih untuk membentuk model klasifikasi yang akan digunakan dalam meramalkan kelas sentimen. Selanjutnya, data uji digunakan sebagai alat evaluasi untuk mengukur seberapa akurat model klasifikasi dalam melakukan klasifikasi.

M. Confusion Matrix

Confusion matrix dilakukan untuk mengevaluasi kinerja model klasifikasi yang digunakan. *Output* dari *confusion matrix* akan disajikan dalam bentuk *heatmap*, dan nilai-nilai kinerja yang diperoleh dari *confusion matrix* meliputi *accuracy*, *precision*, *recall*, dan *f1-score*.

N. Visualisasi Analisis Sentimen

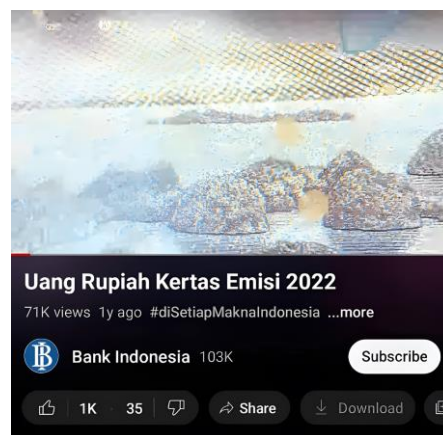
Hasil uji dari analisis sentimen dalam penelitian ini akan disajikan dalam bentuk diagram lingkaran yang menggambarkan hasil sentimen positif dan negatif. Selain itu, akan ditampilkan pula sebuah *wordcloud* yang bertujuan untuk mengilustrasikan kata-kata yang muncul lebih sering dengan ukuran yang lebih besar, sehingga mempermudah observasi secara visual.

4. HASIL DAN PEMBAHASAN

O. Pengumpulan Data

Tahap ini menggunakan YouTube API untuk memperoleh data dari *channel* Bank Indonesia yang berjudul “Uang Rupiah Kertas Emisi 2022” dan video dari *channel* KOMPASTV yang berjudul “Prosesi Peluncuran Tujuh Uang Rupiah Kertas Tahun Emisi 2022, dari Rp 1.000 hingga Rp 100.000”. Data diambil pada tanggal 1 Mei 2023, setelah proses pengambilan data selesai, data dari kedua video tersebut disatukan dengan total 1.259 data, lalu data komentar duplikat dihapus menyisakan total 1237 data, dan kemudian disimpan dalam sebuah *file* dengan format CSV.

Berikut merupakan gambar dari video di *channel* Bank Indonesia yang berjudul “Uang Rupiah Kertas Emisi 2022” dan gambar dari video di *channel* KOMPASTV yang berjudul “Prosesi Peluncuran Tujuh Uang Rupiah Kertas Tahun Emisi 2022, dari Rp 1.000 hingga Rp 100.000” ditunjukkan pada gambar 2 dan gambar 3.



Gambar 2. Video YouTube Channel Bank Indonesia



Gambar 3. Video YouTube Channel KOMPASTV

Berikut merupakan tabel hasil hapus komentar duplikat dari *channel* Bank Indonesia dan dari *channel* KOMPASTV ditunjukkan pada tabel 1.

Tabel 1. Hasil Hapus Komentar Duplikat

No.	Nama	Komentar
1	Santri Kendal Channel	REDENOMINASI RUPIAH + BAHANNYA POLYMER Jadi nggak kalah deh sama uang asing
2	Santri Kendal Channel	Mohon maaf nih pak banyak komentar yang ingin sekali redenominasi rupiah
...
1236	Muhammad Raqwan Rahmatullah	Amiiin
1237	Mammoy	Rubah2 Mulu.,

P. Text Preprocessing

Data yang diperoleh setelah hapus duplikasi data yaitu 1237 data digunakan dalam tahap *text processing* ini dengan urutan, 1.) *case folding*, 2.) *cleaning*, 3.) *Stemming*, 4.) *Tokenizing*. Berikut merupakan tabel hasil *text preprocessing* ditunjukkan pada tabel 2.

Tabel 2. Hasil Text Preprocessing

NO.	Sebelum Text Preprocessing	Sesudah Text Preprocessing
1	REDENOMINASI RUPIAH + BAHANNYA POLYMER Jadi nggak kalah deh sama uang asing	['redenominasi', 'rupiah', 'bahan', 'nya', 'polymerbrjadi', 'nggak', 'kalah', 'deh', 'sama', 'uang', 'asing']
2	Mohon maaf nih pak banyak komentar yang ingin sekali redenominasi rupiah	['mohon', 'maaf', 'nih', 'pak', 'banyak', 'komentar', 'yang', 'ingin', 'sekali', 'redenominasi', 'rupiah']
...
1236	Amiiin	['amiiin']
1237	Rubah2 Mulu.,	['rubah', 'mulu']

Q. Pelabelan

Data *text preprocessing* yang telah diperoleh, kemudian digunakan sebagai data pelabelan. Pelabelan dilakukan secara manual oleh ahli bahasa yaitu guru bahasa Indonesia SMA Islam Bunga Bangsa Samarinda. Data yang dipilih hanya komentar positif dan negatif, kemudian diberi label berupa angka 1 untuk positif dan angka 0 untuk negatif. Data pelabelan yang berhasil dipilih berjumlah 487 komentar yaitu, komentar positif berjumlah 246 dan komentar negatif berjumlah 241. Berikut merupakan tabel data pelabelan ditunjukkan pada tabel 3.

Tabel 3. Data Pelabelan

No.	Komentar	Label
1	['redenominasi', 'rupiah', 'bahan', 'nya', 'polymerbrjadi', 'nggak', 'kalah', 'deh', 'sama', 'uang', 'asing']	1
2	['hahahahaaakalan', 'rezimbilang', 'aja', 'trus', 'terang', 'uang', 'dri', 'rupiah', 'jdi', 'rupiah', 'sy', 'ingat', 'jaman', 'soekarno', 'dulu', 'ya', 'seperti', 'ini', 'cuma', 'beda', 'nya', 'perintah', 'soekarno', 'memberitahu', 'masyarakat', 'dgn', 'buka', 'beda', 'dgn', 'skrg', 'mereka', 'beri', 'alas', 'yg', 'tdk', 'masuk', 'akalberarti', 'apa', 'ya', 'inflasi', 'dan', 'ke', 'ada', 'ekonomi', 'indonesia', 'tdk', 'baik', 'aja', 'kondisi', 'nyasemua', 'jelas', 'tentang', 'uang', 'baru', 'yg', 'nominal', 'jdi', 'jdi', 'dan', 'terus', 'ga', 'masuk', 'akal', 'harus', 'nya', 'perintah', 'cara', 'buka', 'memberitahu', 'masyarakat', 'agar', 'erti', 'dan', 'tau', 'situasi', 'dan', 'kondisi', 'indonesia', 'yg', 'benar', 'nya', 'sy', 'hanya', 'ingat',	0

No.	Komentar	Label
	'kan', 'bagai', 'org', 'tua', 'yg', 'sdh', 'umur', 'hampir', 'perintah', 'jgn', 'trus', 'bohong', 'rakyat', 'nya', 'dgn', 'kalimat', 'dan', 'kata', 'hanya', 'utk', 'tutup', 'yg', 'benar', 'nya']	
...
486	['srimulat', 'spg', 'nya', 'imf']	0
487	['rubah', 'mulu']	0

R. Split Data

Setelah pelabelan, pada tahap ini data dibagi menjadi 2 yaitu data *training* dan data *testing*. Data *training* berjumlah 340 data akan digunakan untuk melatih model, sedangkan data *testing* berjumlah 147 data akan digunakan untuk menguji performa model yang telah dilatih. Berikut merupakan tabel hasil data *training* ditunjukkan pada tabel 4 dan tabel hasil data *testing* ditunjukkan pada tabel 5.

Tabel 4. Hasil Data *Training*

No.	Komentar	Label
1	['mantul', 'href', 'yuk', 'selalu', 'jaga', 'dan', 'rawat', 'rupiah', 'kita', 'dengan', 'terap', 'jangan', 'lipat', 'jangan', 'coret', 'jangan', 'stapler', 'jangan', 'remas', 'dan', 'jangan', 'basah']	1
2	['kenapa', 'sih', 'perintah', 'kok', 'ngutang', 'kan', 'bisa', 'nyetak', 'uang', 'sendiri', 'sekarang', 'cetak', 'uang', 'yg', 'banyak', 'buat', 'bayar', 'utang', 'negara', 'trus', 'jangan', 'ngutang', 'lagi', 'syukur', 'gaji', 'pegawai', 'naik', 'brgitu', 'aja', 'kok', 'repot']	0
...
339	['nilai', 'nominal', 'besar', 'tapi', 'tidak', 'daya']	0
340	['terima', 'kasih', 'banyak', 'buat', 'kontribusi', 'ibu', 'sri', 'mulyani', 'untuk', 'maju', 'indonesia']	1

Tabel 5. Hasil Data *Testing*

No.	Komentar	Label
1	['lebih', 'fresh']	1
2	['redenominasiredenominasiredenominasi', 'ny', 'kebanyak', 'bcotdoang']	0
...
146	['gonta', 'ganti', 'terus', 'gak', 'ada', 'guna', 'nya', 'malah', 'proyek', 'nya', 'bikin', 'kontraktor', 'untung', 'besar', 'yg', 'penting', 'itu', 'nilai', 'mata', 'uang', 'nya', 'ya', 'kan', 'bukan', 'gambar', 'nya']	0
147	['setuju']	1

S. Term Frequency–Inverse Document Frequency (TF-IDF)

Tahap ini menggunakan data pelabelan, kemudian diolah menggunakan TfIdfVectorizer. Bobot tiap kata yang mendekati atau sama dengan nilai 1 menunjukkan bahwa kata tersebut penting dalam kalimat tersebut, sedangkan nilai 0 menunjukkan bahwa kata tersebut tidak ada dalam kalimat atau memiliki frekuensi yang sangat rendah sehingga tidak dianggap penting dalam kalimat tersebut. Berikut merupakan tabel hasil TF-IDF ditunjukkan pada tabel 6.

Tabel 6. Hasil TF-IDF

Komentar	aamin	...	cara	...	deh	...	rubah	...	uang	...	zimbabwe
1	0	...	0	...	0.42	...	0	...	0.13	...	0
2	0	...	0.07	...	0	...	0	...	0.06	...	0
...
486	0	...	0	...	0	...	0	...	0	...	0

Komentar	aamin	...	cara	...	deh	...	rubah	...	uang	...	zimbabwe
487	0	...	0	...	0	...	0.65	...	0	...	0

T. Naïve Bayes Classifier

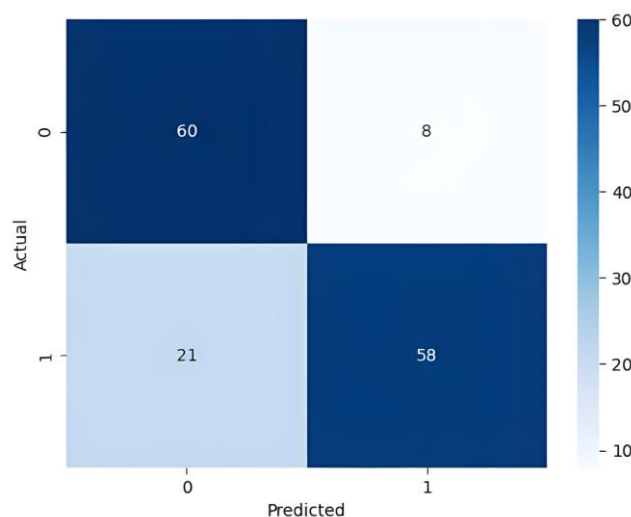
Tahap ini menggunakan data *training* dan data *testing*, kemudian diolah menggunakan MultinomialNB dan CountVectorizer. Model *Naïve Bayes* dilatih terlebih dahulu pada data *training*, setelah itu pelatihan model *Naïve Bayes* digunakan untuk memprediksi hasil pada data *testing* sebagai evaluasi sejauh mana model *Naïve Bayes* dapat melakukan klasifikasi dengan benar terhadap data yang tidak pernah digunakan selama pelatihan, serta menghitung probabilitas hasil prediksi yang akan menjadi label sistem. Label ahli bahasa dilakukan oleh guru bahasa indonesia, sedangkan label sistem dilakukan oleh *Naïve Bayes Classifier*. Berikut merupakan tabel hasil *Naïve Bayes Classifier* ditunjukkan pada tabel 7.

Tabel 7. *Naïve Bayes Classifier*

No.	Komentar	Label Ahli Bahasa	Label Sistem
1	['lebih', 'fresh']	1	1
2	['redenominasiredenominasiredenominasi', 'ny', 'kebanyak', 'bcotdoang']	0	0
...
146	['gonta', 'ganti', 'terus', 'gak', 'ada', 'guna', 'nya', 'malah', 'proyek', 'nya', 'bikin', 'kontraktor', 'untung', 'besar', 'yg', 'penting', 'itu', 'nilai', 'mata', 'uang', 'nya', 'ya', 'kan', 'bukan', 'gambar', 'nya']	0	0
147	['setuju']	1	1

U. Confusion Matrix

Pada tahap ini, evaluasi kinerja *Naïve Bayes Classifier* akan disajikan melalui beberapa metrik, termasuk *heatmap*, *accuracy*, *precision*, *recall*, dan *f1-score*. Metode ini akan memberikan gambaran yang komprehensif tentang sejauh mana model (*Naïve Bayes Classifier*) dapat melakukan klasifikasi dengan benar dan efektif dalam mengidentifikasi kelas target. Berikut merupakan gambar *heatmap confusion matrix* ditunjukkan pada gambar 4 dan berikut merupakan gambar hasil tingkat performa *confusion matrix* ditunjukkan pada gambar 5.



Gambar 4. *Heatmap Confusion Matrix*


```

MultinomialNB Accuracy: 0.8027210884353742
MultinomialNB Precision: 0.8787878787878788
MultinomialNB Recall: 0.7341772151898734
MultinomialNB F1_Score: 0.8
=====
              precision    recall  f1-score   support

         0         0.74         0.88         0.81         68
         1         0.88         0.73         0.80         79

 accuracy                   0.80         147
 macro avg         0.81         0.81         0.80         147
 weighted avg     0.81         0.80         0.80         147
    
```

Gambar 5. Hasil Tingkat Performa *Confusion Matrix*

Penjelasan untuk hasil nilai *heatmap confusion matrix* adalah *True Negative* (TN) berjumlah 60 data, *False Positive* (FP) berjumlah 8 data, *True Positive* (TP) berjumlah 58 data, dan *False Negative* (FN) berjumlah 21 data. Perhitungan manual untuk nilai *accuracy*, *precision*, *recall* dan *f1-score* sebagai berikut:

$$Accuracy = \frac{TN+TP}{TN+FP+FN+TP} = \frac{60+58}{60+8+21+58} = \frac{118}{147} = 0.80 \times 100\% = 80\%$$

$$Precision = \frac{TN}{TN+FP} = \frac{58}{58+8} = \frac{58}{66} = 0.87 \times 100\% = 87\%$$

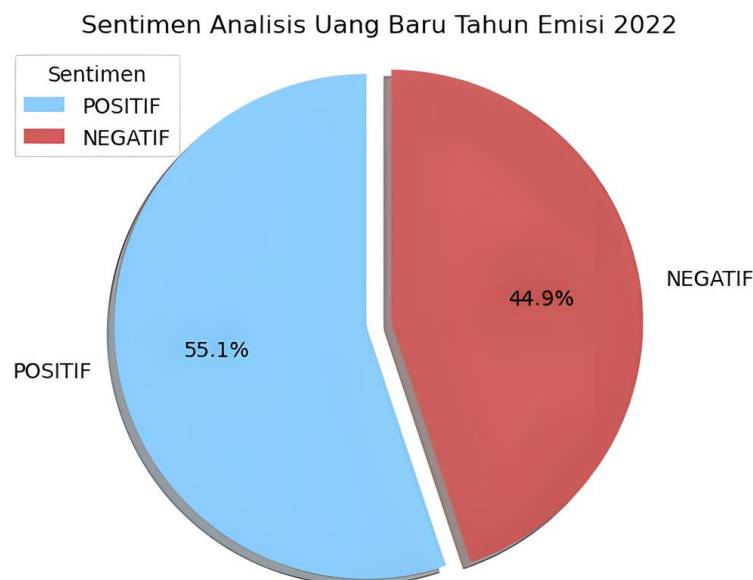
$$Recall = \frac{TP}{TP+FN} = \frac{58}{58+21} = \frac{58}{79} = 0.73 \times 100\% = 73\%$$

$$F1-score = 2 \times \frac{precision \times recall}{precision + recall} = 2 \times \frac{0.87 \times 0.73}{0.87+0.73} = 0.80 \times 100\% = 80\%$$

V. Visualisasi Analisis Sentimen

Terdapat 2 bentuk visualisasi analisis sentimen yaitu diagram lingkaran dan *wordcloud*.

a. Visualisasi diagram lingkaran menampilkan komentar positif berjumlah 82 data (55,1%) berwarna biru dan komentar negatif berjumlah 66 data (44,9%) berwarna merah dari hasil data label sistem (*Naïve Bayes Classifier*).



Gambar 6. Hasil Visualisasi Diagram Lingkaran

b. Visualisasi *wordcloud* menampilkan kata-kata yang muncul dalam teks dengan ukuran kata yang berbeda-beda berdasarkan frekuensi kemunculannya yang terbesar merupakan kalimat yang sering muncul.



Gambar 7. Hasil Visualisasi Wordcloud

5. KESIMPULAN

Berdasarkan hasil penelitian analisis sentimen pengguna YouTube terhadap uang baru tahun emisi 2022 menggunakan metode *Naive Bayes Classifier* dapat ditarik kesimpulan sebagai berikut:

1. Respon dari pengguna YouTube terhadap Uang Baru Tahun Emisi 2022 menunjukkan bahwa sebanyak 55,1% (82 data) termasuk dalam kategori tanggapan positif, sementara 44,9% (66 data) tergolong dalam kategori tanggapan negatif. Analisis ini menggambarkan bahwa mayoritas pengguna YouTube menunjukkan dukungan terhadap Uang Baru Tahun Emisi 2022 berdasarkan tanggapan yang diberikan.
2. Penelitian ini mencapai nilai *accuracy* sebesar 80%, *precision* mencapai 87%, *recall* sebesar 73%, dan *f1-score* sebesar 80%. Angka-angka ini memberikan indikasi bahwa model yang digunakan dalam penelitian ini memiliki kinerja yang baik dan dapat diimplementasikan dengan efektif. Tingkat *accuracy* yang signifikan, *precision* yang tinggi, *recall* yang memadai, dan *f1-score* yang signifikan menyiratkan bahwa model mampu dengan baik mengklasifikasikan sentimen pengguna terhadap Uang Baru Tahun Emisi 2022. Hasilnya, temuan ini memberikan keyakinan bahwa model memiliki kemampuan yang dapat diandalkan untuk analisis sentimen pada respon pengguna YouTube terhadap peristiwa tersebut.

6. DAFTAR PUSTAKA

- Bank Indonesia. (2022). *Bank Indonesia Dan Pemerintah Meluncurkan Uang Rupiah Kertas Tahun Emisi 2022*. https://www.bi.go.id/id/publikasi/ruang-media/news-release/Pages/sp_2421922.aspx
- Chakraborty, K., Bhattacharyya, S., Bag, R., & Hassanien, A. A. (2019). Sentiment Analysis on a Set of Movie Reviews Using Deep Learning Techniques. *Social Network Analytics*, 127–147.
- Hendika Permana, I. P. (2021). Analisis Rasio Pada Akun Youtube Untuk Penelitian Kualitatif Menggunakan Metode Eksploratif. *Jurnal Ilmiah Media Sisfo*, 15(1), 40–48.
- Mas Pintoko, B., & Muslim, K. (2018). Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naive Bayes Classifier. *E-Proceeding of Engineering*, 5(3), 8121–8130.
- Ramadhan, F. A., Sitorus, S. H., & Rismawan, T. (2023). Penerapan Metode Multinomial Naive Bayes untuk Klasifikasi Judul Berita Clickbait dengan Term Frequency - Inverse Document Frequency. *Jurnal Sistem Dan Teknologi Informasi (JustIN)*, 11(1), 70–76. <https://doi.org/10.26418/justin.v11i1.57452>
- Saiyar, H. (2018). Aplikasi Diagnosa Penyakit Tuberculosis Menggunakan Algoritma Naive Bayes. *JURIKOM*, 5(5), 498–502. <http://ejurnal.stmik-budidarma.ac.id/index.php/jurikom|Page|498>
- Suryani, P. S. M., Linawati, L., & Saputra, K. O. (2019). Penggunaan Metode Naive Bayes Classifier pada Analisis Sentimen Facebook Berbahasa Indonesia. *Majalah Ilmiah Teknologi Elektro*, 18(1), 145–148. <https://doi.org/10.24843/mite.2019.v18i01.p22>
- Tutiasri, R. P., Laminto, N. K., & Nazri, K. (2020). Pemanfaatan Youtube Sebagai Media Pembelajaran Bagi Mahasiswa di Tengah Pandemi Covid-19. *Jurnal Komunikasi Masyarakat Dan Keamanan (KOMASKAM)*, 2(2), 1–15.
- Xu, J., Zhang, Y., & Miao, D. (2020). Three-way Confusion Matrix For Classification: A Measure Driven View. *Information Sciences*, 507, 772–794. <https://doi.org/10.1016/J.INS.2019.06.064>
- Yosef, E., Sihombing, S., & Hanggara, B. T. (2020). Pemanfaatan API Youtube dalam Pengembangan Aplikasi Portal Video Penangkaran Kenari untuk Peternak Kenari Berbasis Android. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 4(7), 2067–2074. <http://j-ptiik.ub.ac.id>
- Zainal Macfud, A., Pandu Kusuma, A., Dwi Puspitasari, W., Balitar Blitar JI Majapahit No, I., Sananwetan, K., Blitar, K., & Timur, J. (2023). Analisis Algoritma Naive Bayes Classifier (Nbc) Pada Klasifikasi Tingkat Minat Barang Di Toko Violet Cell. *Jurnal Mahasiswa Teknik Informatika*, 7(1), 87–94.