

---

# Analisis Sentimen Terhadap Layanan Aplikasi Jenius di Media Sosial Menggunakan Algoritma *Long Short-Term Memory*

Modesta Binti Husna<sup>1\*</sup>, Windu Gata<sup>2</sup>

Ilmu Komputer, Universitas Nusa Mandiri, Jakarta, Indonesia

\*e-mail *Corresponding Author*: 14220027@nusamandiri.ac.id

## Abstract

*The Jenius application was launched in 2015 by PT Bank BTPN Tbk, introducing innovation in digital banking services with an active presence on social media platforms like YouTube, Twitter (X), and Instagram. Users often provide reviews that can influence the perception and reputation of the application. This study aims to assist Bank BTPN in enhancing its services by analyzing positive, neutral, and negative user reviews. The method used is the Long Short-Term Memory (LSTM) model for sentiment analysis. The data analyzed consists of 856 comments on YouTube with the keyword "Jenius Application." The LSTM model was implemented to identify general perceptions and aspects that affect user satisfaction, using a data training and testing proportion of 70:30. The study results show that the LSTM model achieved an accuracy of 93.57% on the test dataset, demonstrating the model's effectiveness in identifying user sentiments, thus serving as a tool to improve the quality of Jenius application's services.*

**Keywords:** *Sentiment Analysis; YouTube; Long Short-Term Memory; Jenius App*

## Abstrak

Aplikasi Jenius diluncurkan pada tahun 2015 oleh PT Bank BTPN Tbk, menghadirkan inovasi dalam layanan perbankan digital dengan kehadiran aktif di media sosial seperti *YouTube*, *Twitter (X)*, dan *Instagram*. Pengguna sering memberikan ulasan yang dapat mempengaruhi persepsi dan reputasi aplikasi tersebut. Penelitian ini bertujuan untuk membantu Bank BTPN meningkatkan layanan dengan menganalisis ulasan positif, netral, dan negatif dari pengguna. Metode yang digunakan adalah model *Long Short-Term Memory (LSTM)* untuk analisis sentimen. Data yang dianalisis terdiri dari 856 komentar di *YouTube* dengan kata kunci "Aplikasi Jenius." Model LSTM diimplementasikan untuk mengidentifikasi persepsi umum dan aspek-aspek yang mempengaruhi kepuasan pengguna, dengan parameter proporsi data latih dan data uji sebesar 70:30. Hasil penelitian menunjukkan bahwa model LSTM mencapai akurasi sebesar 93.57% pada dataset uji, menunjukkan efektivitas model dalam mengidentifikasi sentimen pengguna sehingga dapat digunakan sebagai alat untuk meningkatkan kualitas layanan aplikasi Jenius.

**Kata kunci:** *Analisis Sentimen; YouTube; Long Short-Term Memory; Aplikasi Jenius*

## 1. Pendahuluan

Pada Masa digital saat ini, media sosial telah menjadi tempat utama bagi masyarakat untuk menyampaikan opini atau pendapat dan pengalaman mereka terkait berbagai produk dan layanan. Perbankan, sebagai salah satu sektor yang intensif berinteraksi dengan pelanggan, tidak terkecuali dari fenomena ini. Aplikasi Jenius sebagai inovator dalam layanan perbankan digital di Indonesia, sering menjadi subjek dari berbagai diskusi di platform media sosial maupun di lingkungan masyarakat. Analisis sentimen dari data tersebut dapat memberikan wawasan berharga mengenai persepsi pelanggan dan membantu dalam pengambilan keputusan strategis oleh manajemen bank[1].

Aplikasi Jenius, pertama kali di kenalkan oleh PT Bank Tabungan Pensiunan Nasional Tbk (BTPN) pada tahun 2016, merupakan inovator dalam layanan perbankan digital di Indonesia. Meskipun telah memiliki fitur yang fleksibel, intuitif, dan mudah diakses melalui teknologi digital, ulasan pengguna di media sosial menunjukkan adanya perbedaan antara harapan ideal dan kondisi nyata yang dialami pengguna. Pengguna sering menyampaikan

pendapat mereka melalui komentar di *YouTube*, dan analisis terhadap komentar-komentar ini menunjukkan adanya sentimen beragam yang mencerminkan kepuasan maupun ketidakpuasan terhadap layanan yang diberikan[2].

Penelitian ini mengajukan penggunaan model Long Short-Term Memory (LSTM), yang merupakan varian dari jaringan saraf tiruan yang efektif untuk menangani data sekuensial seperti teks, dalam klasifikasi sentimen dari komentar dan ulasan pengguna tentang layanan Aplikasi Jenius di media sosial. Dengan menerapkan LSTM, penelitian ini bertujuan untuk mengidentifikasi dan mengelompokkan opini pelanggan menjadi sentimen positif, negatif, dan netral. Pendekatan ini didukung oleh beberapa penelitian terdahulu yang menunjukkan efektivitas LSTM dalam analisis sentimen (Graves, 2012; Hochreiter & Schmidhuber, 1997) [3] serta penggunaan algoritma *machine learning* seperti *SVM*, *Naive Bayes*, *KNN*, *Decision Tree*, dan *Logistic Regression* dalam konteks serupa.

Tujuan utama dari penelitian ini adalah untuk mengidentifikasi dan mengelompokkan opini pelanggan terhadap Aplikasi Jenius, memberikan gambaran lebih jelas mengenai kinerja layanan bank secara *real-time*. Penelitian ini juga bertujuan mengevaluasi efektivitas LSTM dalam konteks analisis sentimen serta mengidentifikasi faktor-faktor yang paling mempengaruhi sentimen pelanggan. Manfaat yang diharapkan dari penelitian ini adalah memberikan tolok ukur bagi Bank BTPN untuk meningkatkan layanan berdasarkan ulasan positif dan negatif, sehingga dapat meningkatkan kepuasan dan loyalitas pelanggan secara keseluruhan.

## 2. Tinjauan Pustaka

Berbagai penelitian telah dilakukan mengenai analisis sentimen, di mana para peneliti telah menerapkan berbagai metode klasifikasi untuk mengatasi masalah ini. Berikut adalah rangkuman dari beberapa penelitian tersebut:

Penelitian ini bertujuan untuk menganalisis komentar atau opini terhadap layanan Aplikasi Jenius di berbagai media sosial khususnya *Youtube* dengan menerapkan algoritma *Long Short-Term Memory* (LSTM). Metodologi ini dipilih karena efektivitasnya dalam menangani data teks sekuensial, seperti yang telah dibuktikan dalam berbagai studi sebelumnya. Misalnya, Karthik Gopalakrishnan dan Fathi M. Salem menunjukkan bahwa metodologi LSTM sangat efisien dalam analisis sentimen, memungkinkan pemrosesan data teks dengan akurasi tinggi.[4]

Dalam konteks perbankan, penggunaan LSTM untuk analisis sentimen telah diterapkan secara luas. Studi oleh Nurrohmat dan Noor membahas penggunaan LSTM dalam analisis ulasan novel, yang menunjukkan bahwa LSTM mampu mengklasifikasikan sentimen dengan lebih akurat dibandingkan metode lain seperti *Naive Bayes*. Model *deep learning* LSTM yang mereka kembangkan berhasil menganalisis dan mengkategorikan keluhan ke dalam departemen terkait dengan akurasi sebesar 84%. Otomatisasi ini secara signifikan mengurangi beban kerja manual dan meningkatkan efisiensi dalam menangani keluhan pelanggan M. Amin Nurrohmat and A. S. Noor [2]. Selain itu, proyek yang dilakukan oleh Sinha menunjukkan bahwa algoritma LSTM dapat digunakan untuk menganalisis dan mengkategorikan keluhan nasabah bank secara efektif, memberikan wawasan berharga bagi peningkatan layanan pelanggan [5].

Penelitian oleh Ozbayoglu, Gudelek, dan Sezer juga mendukung penggunaan LSTM dalam aplikasi keuangan, menunjukkan bahwa model *deep learning* ini dapat diterapkan untuk memprediksi pergerakan pasar saham berdasarkan analisis berita keuangan [6]. Hal ini relevan dalam konteks perbankan karena menunjukkan kemampuan LSTM untuk menganalisis data teks yang kompleks dan menghasilkan prediksi yang akurat. Studi lain oleh Li, Xie, Chen, Wang, dan Deng menguraikan bagaimana analisis sentimen dapat digunakan untuk memahami dampak berita terhadap pergerakan harga saham, yang dapat diterapkan untuk menganalisis sentimen dalam perbankan [7].

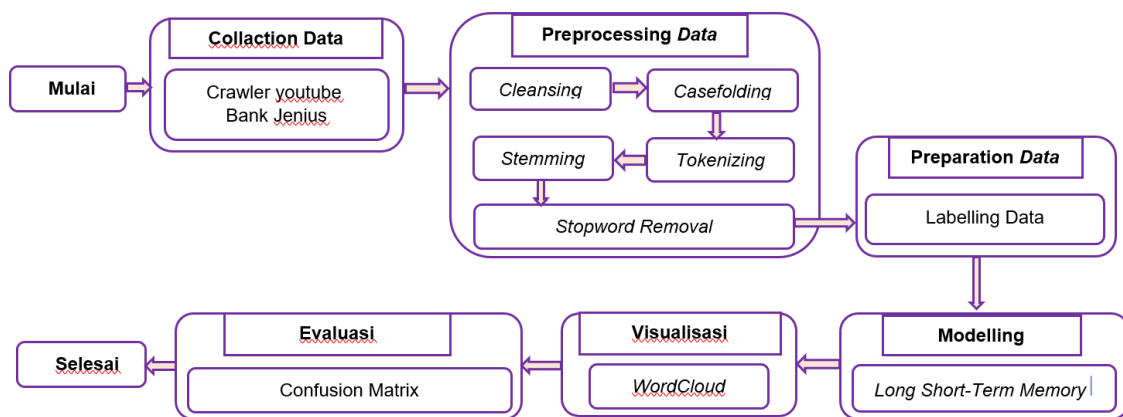
Ray dan Chakrabati mengusulkan pendekatan gabungan antara metode berbasis aturan dan pembelajaran mendalam untuk analisis sentimen tingkat aspek, menggunakan jaringan neural konvolusional tujuh lapis untuk menandai setiap aspek dalam komentar, yang juga dapat diterapkan dalam penelitian ini untuk menganalisis aspek-aspek spesifik dari layanan Jenius[8].

Penelitian ini mengusulkan analisis sentimen terhadap layanan aplikasi Jenius di media sosial menggunakan algoritma *Long Short-Term Memory* (LSTM), yang merupakan peningkatan signifikan dibandingkan penelitian terdahulu yang umumnya implementasikan metode seperti Klasifikasi (*Naive Bayes*), *SVM*, Pohon Keputusan, dan *Logistic Regression*.

Keunikan studi ini terletak pada penerapan LSTM yang lebih efektif dalam mengelola data teks sekuensial dari *YouTube*, serta penggunaan teknik *preprocessing* yang menyeluruh seperti pembersihan data, case folding, tokenisasi, stemming, dan penghapusan kata-kata umum. Selain itu, studi ini juga mengatasi ketidakseimbangan kelas dengan menggunakan teknik resampling dan pembobotan kelas. Dengan data yang dianalisis sebanyak 856 komentar, penelitian ini mencapai akurasi tinggi sebesar 93.57%, menunjukkan efektivitas LSTM dalam mengidentifikasi sentimen pengguna dan menawarkan potensi implementasi *real-time* untuk meningkatkan kualitas layanan Bank BTPN berdasarkan analisis sentimen yang diperoleh.

### 3. Metodologi

Dalam penelitian ini, metode analisis yang diterapkan adalah LSTM (*Long Short-Term Memory*) untuk mengoptimalkan hasil. Prosesnya terdiri dari beberapa tahapan penting, yaitu pengumpulan data, pelabelan data, dan preprocessing. Berikut adalah uraian lengkap dari masing-masing tahapan tersebut.



Gambar 1. Metode Penelitian

#### 3.1 Crawling Data

Pada langkah awal penelitian, peneliti menggunakan query untuk video *YouTube* yang berkaitan dengan Aplikasi Jenius sebagai sumber kumpulan data. Data ini diperoleh dari layanan berbagi video *YouTube*. Data dikumpulkan melalui teknik *crawling* dengan menggunakan *Application Programming Interface* (API) atau menggunakan bahasa pemrograman Python yang dimodifikasi dengan memanfaatkan *library Selenium Webdriver*.

#### 3.2 Data Preprocessing

Data *preprocessing* adalah metode yang digunakan untuk mengubah data, termasuk validasi dan imputasi untuk mengetahui seberapa lengkap dan akurat data yang akan digunakan dalam proses analisis selanjutnya. Tahapan *preprocessing* termasuk pembersihan, penyusunan case, tokenisasi, stemming, dan penghapusan kata-kata umum.[9].

1) *Clean sing*

Proses pembersihan kata-kata yang tidak berguna dari data untuk menghilangkan kebisingan. Pembersihan data adalah langkah kritis dalam pemrosesan bahasa alami dan proyek machine learning, yang melibatkan penghapusan atau perbaikan data yang salah atau tidak akurat untuk meningkatkan kinerja model. Teknik ini mencakup metode manual dan otomatis, dengan solusi seperti *AlphaClean* dan *CPClean* yang mengotomatiskan proses ini untuk memastikan efisiensi dan cakupan optimal. Dengan menggunakan metode ini, data yang bersih dapat diperoleh, yang pada akhirnya meningkatkan akurasi dan keandalan analisis data dan model *machine learning* [10].

2) *Case Folding*

Merupakan proses penyalarsan teks dalam kumpulan data. Dalam penelitian ini, data akan diubah menjadi huruf kecil melalui proses *Case Folding*. [11].

3) *Removal Symbol*

Ini adalah proses menghilangkan *prefix* atau simbol yang tidak berguna dari kalimat, seperti @ yang menunjukkan pengguna atau tautan http://.

4) *Tokenizing*

Tahap memecah atau memisahkan sebuah kalimat menjadi bagian-bagian yang lebih. Pemecahan ini dilakukan dengan tanda spasi sebagai pemisah antar kata. Dalam penelitian yang dipaparkan oleh Jurafsky dan Martin (2021) dalam buku "*Speech and Language Processing*" edisi ketiga, dijelaskan bahwa tokenisasi merupakan langkah awal yang krusial dalam pemrosesan bahasa alami. Penelitian ini menyoroti berbagai teknik tokenisasi, termasuk tokenisasi berbasis spasi, karakter, kata, dan sub-kata. Jurafsky dan Martin menekankan pentingnya tokenisasi yang tepat dalam meningkatkan akurasi analisis linguistik lebih lanjut, seperti dalam tugas-tugas klasifikasi teks dan pemodelan bahasa.[12]

5) *Stemming*

Proses mengembalikan kata ke dalam kata dasarnya atau bentuk standar[13]. *Stemming* adalah metode dalam pemrosesan bahasa alami (Natural Language Processing/NLP) yang mengurangi kata-kata ke bentuk dasarnya dengan menghilangkan akhiran. Teknik ini esensial dalam berbagai tugas NLP seperti pengambilan informasi, klasifikasi teks, dan analisis sentimen. Dengan menyederhanakan kata-kata ke bentuk dasarnya, *Stemming* mengurangi keragaman kata dan meningkatkan efisiensi analisis data. Algoritma *Stemming* yang umum termasuk Porter, Snowball, dan Lovins, yang masing-masing menawarkan kecepatan dan akurasi dalam berbagai konteks aplikasi [14].

6) *Stopword Removal*

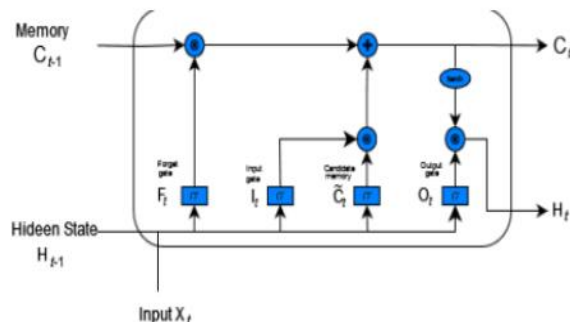
Menghapus kata-kata yang dianggap tidak terlalu penting. *Stopword removal* adalah langkah krusial dalam pemrosesan bahasa alami (*Natural Language Processing/NLP*) yang bertujuan untuk menghilangkan kata-kata umum yang tidak memberikan informasi signifikan dalam analisis teks. Kata-kata seperti "dan", "yang", "untuk" sering muncul dalam teks tetapi tidak memiliki nilai analitis yang tinggi. Dengan menghapus *stopword*, algoritma NLP dapat lebih fokus pada kata-kata yang lebih bermakna dan relevan untuk tugas-tugas seperti klasifikasi teks, pengambilan informasi, dan pemodelan topik. [15]

### 3.3 Analisa atau Labeling Data

Proses Analisis atau Pelabelan adalah upaya untuk menentukan kategori pada sebuah kumpulan data atau dataset, seperti komentar dari *YouTube*, ke dalam beberapa kategori yang diinginkan, yaitu positif, netral, atau negatif. Pelabelan ini bergantung pada nilai dari susunan kalimat yang ada dalam dataset [16]. Kumpulan data atau dataset akan dianotasi menjadi kelas sentimen positif dan negatif melalui tiga metode, yaitu VADER, *Text Blob*, dan manual.

### 3.4 Modelling

Banyak algoritma yang digunakan dalam penelitian tentang analisis sentimen. Penelitian ini akan menggunakan algoritma *Long-Short Term Memory* (LSTM). Algoritma ini unggul dalam membangun model prediksi dan merupakan turunan dari *Recurrent Neural Network* (RNN) yang dirancang untuk mengolah data berurutan. LSTM dibuat untuk mengatasi masalah gradien pada RNN, terutama ketika menghadapi masalah *vanishing* dan *exploding gradient*[17].



Gambar 2. Model LSTM

### 3.5 WordCloud

WordCloud digunakan untuk menampilkan kata-kata yang digunakan dalam penelitian ini. Semakin sering sebuah kata muncul, semakin besar tampilannya dalam grafik kata. Dengan *WordCloud*, teks yang paling sering muncul dalam dataset dapat direpresentasikan secara visual.[18].

### 3.6 Evaluasi

Proses Untuk mengetahui seberapa baik kinerja suatu model, proses prediksi dapat diukur. Dalam hal ini, akurasi untuk seluruh dataset diperoleh selama pelatihan. Untuk menghitung akurasi, Persamaan 1 menggunakan tiga kategori: positif, netral, dan negatif. TP adalah Nilai Positif Asli, TNt adalah Nilai Neutral Asli, dan TNg adalah Nilai Negatif Asli. FP1 adalah Nilai Positif Palsu 1, FP2 adalah Nilai Positif Palsu 2, FNg1 adalah Nilai Negatif Palsu 1, dan FNt2 adalah Nilai Negatif Palsu 2.

$$Accuracy = \frac{TP+TNt+TNg}{TP+FP1+FP2 +TNt+FNt1+FNt2+TNg+TNg1+TNg2} \tag{1}$$

Persamaan 2 digunakan untuk menghitung kemungkinan kelas yang memiliki nilai benar yang positif atau benar. Persamaan 3 dan 4 juga digunakan untuk menghitung hasil rata-rata makro (*macro avg*) dan rata-rata berbobot (*weighted avg*).

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

$$Macro\ Av = \frac{\sum_{i=1}^n \frac{TP}{TP+FN}}{l} \tag{3}$$

$$Weighted\ Av = \frac{\sum_{i=1}^n \frac{TP}{TP+FN} * ni}{l} \tag{4}$$

Presisi, juga dikenal sebagai *precision*, adalah proses untuk mengukur keakuratan prediksi pada nilai benar, terlepas dari fakta bahwa data yang digunakan sudah termasuk dalam kelas yang benar atau positif. Untuk menghitung ketepatan, persamaan 5 diberikan.

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

Seperti yang tercantum dalam persamaan 6, *F1-Score* adalah metrik evaluasi yang digunakan untuk menilai keseimbangan antara *precision* dan *recall*. *F1-Score* menunjukkan seberapa optimal kinerja model dengan menggabungkan nilai *precision* dan *recall*. Persamaan yang digunakan untuk menghitung *F1-Score* adalah sebagai berikut:

$$F1 - score = \frac{2 * Recall * Precisio}{Recall + Precisio} \tag{6}$$

Dalam prediksi klasifikasi aktual, persamaan 7, 8, dan 9 adalah metrik evaluasi yang digunakan untuk mengukur seberapa baik model klasifikasi berfungsi dalam mengidentifikasi kelas pada berbagai tingkat ambang batas (*threshold*), tanpa mempertimbangkan pilihan ambang batas tertentu.

$$TPR = \frac{TP}{TP+FN1} \tag{7}$$

$$FPR = \frac{FP1}{TN+TN1+TN2} \tag{8}$$

$$AUC = \frac{TPR+(1-FPR)}{2} \tag{9}$$

#### 4. Hasil dan Pembahasan

Penelitian yang melakukan analisis sentimen komentar pengguna layanan Aplikasi Jenius di media sosial (menggunakan metode *Long Short-Term Memory*, atau LSTM) akan menjadi dasar perancangan sistem analisis sentimen, kami berhasil mengidentifikasi pola-pola sentimen dari ratusan komentar yang dianalisis. Pembahasan ini akan menguraikan secara mendetail distribusi sentimen yang ada, serta implikasi dari temuan ini terhadap pemahaman kita mengenai persepsi konsumen terhadap layanan Aplikasi Jenius."

##### 4.1 Crawling Data

Data yang berhasil dikumpulkan dari *YouTube* sebanyak 918 komentar, diambil dari beberapa video *YouTube* yang berhubungan dengan komentar pelanggan dan calon pelanggan Aplikasi Jenius, seperti yang terlihat pada gambar 3.

	tanggal	nama	komentar	jumlah_like
0	2022-07-19T13:11:54Z	@DeraHidayat	Review Kartu ATM Jenius, Cara dapat Kartunya 🙌 ...	1
1	2024-05-02T01:14:46Z	@delladhafin8464	pakai akun dana nggak ada potongan	0
2	2024-04-12T01:51:35Z	@AgusTina-hk6vi	Aku nasabah Taspen Mandiri , bisa pinjam di Je...	0
3	2024-04-08T07:50:54Z	@JamirsyahParulianSinaga-nm8st	aplikasi abal-abal yg gak mutu. Pret dengan...	0
4	2024-04-04T08:26:16Z	@imammuddin4622	Klo dpt Rp 5.000.000 tarik hrs transfer Rp 73...	1
...	...	...	...	...
128	2022-05-30T06:58:34Z	@habibahputra1685	Bg mau tanya seputaran akulaku. Tgggl 28 kmaren...	0
129	2022-05-30T15:02:50Z	@DeraHidayat	Ekspedisi nya apa emang	1
130	2022-06-16T19:58:14Z	@habibahputra1685	@@DeraHidayat uda kak aman. 🤔 cuma mau tanya k...	0
131	2022-06-18T08:54:10Z	@DeraHidayat	Aman sih, asal kode pembayaran nya resmi dari ...	0
132	2022-07-23T13:58:26Z	@rajasihombing2627	@@habibahputra1685 salah kamar nih orang 🤔	0

Gambar 3. Hasil *Crawling* Data

##### 4.2 Preprocessing Data

Pra-pemrosesan teks, atau *text preprocessing*, adalah proses membersihkan data sebelum diolah. Tahap ini meliputi lima langkah: *cleansing* (membersihkan data dari karakter atau simbol yang tidak penting), *case folding* (mengubah semua huruf menjadi huruf kecil), *tokenizing* (memecah teks menjadi kata-kata), *stemming* (mengembalikan kata ke bentuk dasarnya), dan *stopword removal* (menghapus kata-kata umum yang tidak penting). Berikut proses dari tahapan tersebut:

###### 1) *Clean* sing

	komentar	clean
0	Bank digital itu aman gak sih kalo kita mau de...	Bank digital itu aman gak sih kalo kita mau de...
1	Kak cara isi saldo nya gmna	Kak cara isi saldo nya gmna
2	Kak mau tanya..aku punya credit di jenius tapi...	Kak mau tanyaaku punya credit di jenius tapi s...
3	Hati hati lebih baik jangan download jenius..u...	Hati hati lebih baik jangan download jeniusuda...
4	Jenius Bisa req kartu kredit kak	Jenius Bisa req ka u kredit kak
...	...	...
851	Bg mau tanya seputaran akulaku. Tgggl 28 kmaren...	Bg mau tanya seputaran akulaku Tgggl kmaren pe...
852	Ekspedisi nya apa eman	Ekspedisi nya apa eman
853	@@DeraHidayat uda kak aman. 🤔👉 cuma mau tany...	DeraHidayat uda kak aman cuma mau tanya kak ka...
854	Aman sih; asal kode pembayaran nya resmi dari ...	Aman sih asal kode pembayaran nya resmi dari a...
855	@@habibahputra1685 salah kamar nih orang 🤔	habibahputra salah kamar nih orang

Gambar 4. Hasil Komentar Setelah Proses *Clean* sing

Pada gambar 4. tersebut menampilkan hasil proses pembersihan (*clean sing*) dari data komentar pengguna terkait aplikasi perbankan digital Jenius. Pada kolom "komentar," terdapat komentar asli dari pengguna, sementara kolom "*clean* " menunjukkan versi komentar yang telah dibersihkan dari elemen-elemen tidak relevan seperti emoji, simbol, dan kata-kata tidak penting. Misalnya, komentar pada baris pertama "Bank digital itu aman gak sih kalo kita mau de..." telah dibersihkan menjadi "Bank digital itu aman gak sih kalo kita mau de...". Proses pembersihan ini penting untuk memastikan analisis sentimen yang dilakukan kemudian hanya berdasarkan teks yang relevan dan bersih, sehingga hasil analisis lebih akurat dan dapat diandalkan. Langkah ini adalah tahap awal yang sangat penting dalam analisis teks untuk menghilangkan noise yang bisa mengganggu analisis lebih lanjut.

2) *Case Folding*

	<i>clean</i>	<i>lowertext</i>
0	Bank digital itu aman gak sih kalo kita mau de...	bank digital itu aman gak sih kalo kita mau de...
1	Kak cara isi saldo nya gmna	kak cara isi saldo nya gmna
2	Kak mau tanyaaku punya credit di jenius tapi s...	kak mau tanyaaku punya credit di jenius tapi s...
3	Hati hati lebih baik jangan download jeniusuda...	hati hati lebih baik jangan download jeniusuda...
4	Jenius Bisa req ka u kredit kak	jenius bisa req ka u kredit kak
5	Saya pengguna Jenius berharap semoga ke depann...	saya pengguna jenius berharap semoga ke depann...
6	Hallo kak q udah pakek ATM sering tarik tunai ...	hallo kak q udah pakek atm sering tarik tunai ...
7	Apakah bisa untuk membayar E Visa Turkey yg me...	apakah bisa untuk membayar e visa turkey yg me...
8	Req transfer dari jenius ke bank luar neger	req transfer dari jenius ke bank luar neger
9	Kak bgmn cara request ka unyasoalnya sdh aktif...	kak bgmn cara request ka unyasoalnya sdh aktif...

Gambar 5. Hasil Komentar Setelah Proses Case Folding

Gambar 6. Menunjukkan Proses *case folding* di gunakan untuk menyamakan format teks dengan mengubah seluruh kalimat menjadi huruf kecil, sehingga menghilangkan perbedaan antara huruf besar dan kecil yang bisa mempengaruhi analisis lebih lanjut. Sebagai contoh, komentar "Bank digital itu aman gak sih kalo kita mau de..." pada kolom "*clean* " diubah menjadi "bank digital itu aman gak sih kalo kita mau de..." pada kolom "*lowertext*". Langkah ini penting dalam pemrosesan teks karena membantu mengurangi variasi yang tidak perlu dalam data, sehingga mempermudah tahap analisis sentimen dan pengolahan bahasa alami selanjutnya.

3) *Tokenizing*

Tabel 1. menampilkan hasil proses "*Tokenizing*" pada data komentar pengguna aplikasi Jenius. Kolom "*clean* " memuat komentar yang telah dibersihkan dari komponen-komponen tidak relevan, sementara kolom "*Tokenizing*" menunjukkan hasil setelah komentar tersebut dipecah menjadi token-token individual. Proses *Tokenizing* ini bertujuan untuk memecah teks menjadi unit-unit kata yang lebih kecil sehingga lebih mudah dianalisis secara mendalam. Sebagai contoh, komentar "Bank digital itu aman gak sih kalo kita mau deposito diatas jt" dipecah menjadi token-token seperti ['bank', 'digital', 'itu', 'aman', 'gak', 'sih', 'kalo', 'kita', 'mau', 'deposito', 'didas', 'jt']. Langkah ini penting dalam pemrosesan teks karena memungkinkan analisis yang lebih rinci dan akurat pada level kata.

Tabel 1. Hasil Komentar setelah Proses *Tokenizing*

	clean	tokenizing
1	Bank digital itu aman gak sih kalo kita mau deposito diatas jt	['bank', 'digital', 'itu', 'aman', 'gak', 'sih', 'kalo', 'kita', 'mau', 'deposito', 'diatas', 'jt']
2	Kak cara isi saldo nya gmna	['kak', 'cara', 'isi', 'saldo', 'nya', 'gmna']
3	Kak mau tanyaaku punya credit di jenius tapi sudah lunas dan aku mau sambung lgi credit aku trus cara nya gimana kak	['kak', 'mau', 'tanyaaku', 'punya', 'credit', 'di', 'jenius', 'tapi', 'sudah', 'lunas', 'dan', 'aku', 'mau', 'sambung', 'lgi', 'credit', 'aku', 'trus', 'cara', 'nya', 'gimana', 'kak']
4	Hati hati lebih baik jangan download jeniusudah daftar kemudian isi data Dan verifikasi langsung gak aktifcuman ngambil data doang	['hati', 'hati', 'lebih', 'baik', 'jangan', 'download', 'jeniusudah', 'daftar', 'kemudian', 'isi', 'data', 'dan', 'verifikasilangsung', 'gak', 'aktifcuman', 'ngambil', 'data', 'doang']
5	Jenius Bisa req ka u kredit kak	['jenius', 'bisa', 'req', 'ka', 'u', 'kredit', 'kak']
6	Saya pengguna Jenius berharap semoga ke depannya ada Jenius syariah	['saya', 'pengguna', 'jenius', 'berharap', 'semoga', 'ke', 'depannya', 'ada', 'jenius', 'syariah']

#### 4) Stemming

proses *Stemming* pada tabel 2. Merupakan data komentar pengguna aplikasi Jenius. Kolom "*clean*" berisi komentar yang telah dibersihkan dari komponen-komponen tidak relevan, sementara kolom "*Stemming*" menunjukkan hasil setelah kata-kata dalam komentar diubah menjadi bentuk dasar atau akarnya. Proses *Stemming* ini bertujuan untuk menghilangkan imbuhan pada kata sehingga setiap kata dapat dianalisis dalam bentuk dasarnya, yang membantu dalam konsistensi analisis teks. Misalnya, komentar "Bank digital itu aman gak sih kalo kita mau deposito diatas jt" diubah menjadi token-token dasar seperti ['bank', 'digital', 'itu', 'aman', 'gak', 'sih', 'kalo', 'kita', 'mau', 'deposito', 'atas', 'jt']. Langkah ini penting dalam pemrosesan teks karena memungkinkan analisis yang lebih akurat dan efektif pada level kata dasar, yang sangat berguna dalam aplikasi seperti analisis sentimen, klasifikasi teks, dan aplikasi lain dalam pemrosesan bahasa alami. Dengan *Stemming*, kita dapat mengurangi variasi kata yang disebabkan oleh bentuk infleksi yang berbeda, sehingga meningkatkan akurasi dan relevansi hasil analisis teks.

Tabel 2. Hasil Komentar setelah Proses *Stemming*

	clean	stemming
1	Bank digital itu aman gak sih kalo kita mau deposito diatas jt	['bank', 'digital', 'itu', 'aman', 'gak', 'sih', 'kalo', 'kita', 'mau', 'deposito', 'atas', 'jt']
2	Kak cara isi saldo nya gmna	['kak', 'cara', 'isi', 'saldo', 'nya', 'gmna']
3	Kak mau tanyaaku punya credit di jenius tapi sudah lunas dan aku mau sambung lgi credit aku trus cara nya gimana kak	['kak', 'mau', 'tanyaaku', 'punya', 'credit', 'di', 'jenius', 'tapi', 'sudah', 'lunas', 'dan', 'aku', 'mau', 'sambung', 'lgi', 'credit', 'aku', 'trus', 'cara', 'nya', 'gimana', 'kak']
4	Hati hati lebih baik jangan download jeniusudah daftar kemudian isi data Dan verifikasi langsung gak aktifcuman ngambil data doang	['hati', 'hati', 'lebih', 'baik', 'jangan', 'download', 'jeniusudah', 'daftar', 'kemudian', 'isi', 'data', 'dan', 'verifikasilangsung', 'gak', 'aktifcuman', 'ngambil', 'data', 'doang']
5	Jenius Bisa req ka u kredit kak	['jenius', 'bisa', 'req', 'ka', 'u', 'kredit', 'kak']
6	Saya pengguna Jenius berharap semoga ke depannya ada Jenius syariah	['saya', 'guna', 'jenius', 'harap', 'moga', 'ke', 'depan', 'ada', 'jenius', 'syariah']

5) *Stopword Removal*

*Stopword Removal* yang ditampilkan pada tabel 3. merupakan hasil proses penghapusan kata-kata umum (*stopwords*) pada data komentar pengguna aplikasi Jenius. Kolom "*clean*" berisi komentar yang telah dibersihkan dari komponen-komponen tidak relevan, sedangkan kolom "*stopword*" menunjukkan hasil setelah kata-kata umum yang tidak memberikan banyak nilai analisis, seperti "itu", "sih", "dan", dihapus dari komentar. Sebagai contoh, komentar "Bank digital itu aman gak sih kalo kita mau deposito diatas jt" diubah menjadi "bank digital aman gak kalo kita mau deposito atas jt". Proses ini penting dalam pemrosesan teks karena menghilangkan kata-kata yang sering muncul tetapi tidak memiliki makna signifikan, sehingga membantu meningkatkan fokus pada kata-kata kunci yang lebih relevan dalam analisis sentimen. Dengan menghapus *stopwords*, analisis teks dapat menjadi lebih efisien dan akurat, memungkinkan untuk mengidentifikasi pola dan tren yang lebih jelas dalam data komentar pengguna.

Tabel 3. Hasil Komentar setelah Proses Stopword Removal

	clean	stopword
1	Bank digital itu aman gak sih kalo kita mau deposito diatas jt	bank digital itu aman gak sih kalo kita mau deposito atas jt
2	Kak cara isi saldo nya gmna	kak cara isi saldo nya gmna
3	Kak mau tanyaaku punya credit di jenius tapi sudah lunas dan aku mau sambung lgi credit aku trus cara nya gimana kak	kak mau tanyaaku punya credit di jenius tapi sudah lunas dan aku mau sambung lgi credit aku trus cara nya gimana kak
4	Hati hati lebih baik jangan download jeniusudah daftar kemudian isi data Dan verifikasi langsung gak aktifcuman ngambil data doang	hati hati lebih baik jangan download jeniusudah daftar kemudian isi data dan verifikasi langsung gak aktifcuman ngambil data doang
5	Jenius Bisa req ka u kredit kak	jenius bisa req ka u kredit kak
6	Saya pengguna Jenius berharap semoga ke depannya ada Jenius syariah	saya guna jenius harap moga ke depan ada jenius syariah
7	Hallo kak q udah pakek ATM sering tarik tunai di agen BRI Links brApakah jenius bisa tarik tunai di agen BRI Links kah	hallo kak q udah pakek atm sering tarik tunai di agen bri links brapakah jenius bisa tarik tunai di agen bri links kah
8	Apakah bisa untuk membayar E Visa Turkey yg menggunakan CC	apakah bisa untuk bayar e visa turkey yg guna cc
9	Req transfer dari jenius ke bank luar negeri	req transfer dari jenius ke bank luar negeri

4.3 *Labelling Data*

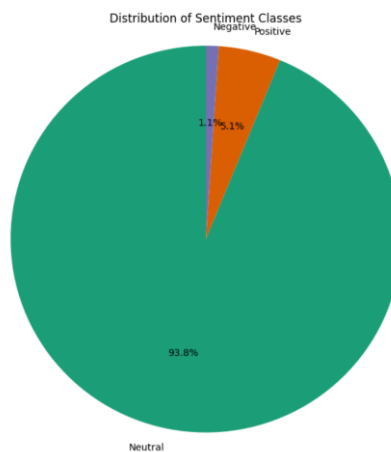
Proses pelabelan dilakukan untuk memberikan label pada komentar yang telah dikumpulkan dan disimpan pada komentar\_youtube.csv. Pada tahap sebelumnya diawal telah dilakukan proses *Cleansing, Tokenizing, Case Folding, Stemming* dan *Penghapusan Stopwords*. Setelah proses-proses tersebut selanjutnya dilakukan Labeling Sentimen sebagai langkah terakhir. Pelabelan sentimen dilakukan menggunakan beberapa metode: VADER (*Valence Aware Dictionary and sEntiment Reasoner*), alat berbasis leksikon yang dirancang khusus untuk analisis sentimen pada teks media sosial; *Text Blob*, pustaka Python yang menyediakan alat untuk analisis sentimen dan klasifikasi teks; serta pelabelan manual yang dilakukan oleh manusia dengan memahami konteks dan nuansa teks. Dengan memanfaatkan ketiga metode ini, setiap komentar atau teks dianalisis dan diberi label sebagai sentimen positif, negatif, atau netral sebagaimana terlihat pada tabel 3. Proses ini memastikan bahwa hasil analisis sentimen akurat dan mempertimbangkan berbagai perspektif, baik dari alat otomatis maupun dari evaluasi manusia.

Tabel 3. Hasil Labeling Data

text	Sentiment
mbak nya pinter bange	Neutral
untuk top up tabungan jenius dari rekening lain kena biaya admin juga akhirnya	Positive
minimal jt mbakbrbuat ngecek lengkapnya dari awards aj	Negative
berlianafebriyanti saldo rata-rata selama bulan sebelumnya minimal rpjuta infonya ada di menu awards mbak	Negative
ok juga pembobol banknya ya mengapa pula nabung jenius jika hanya menguntungkan pembobol ban	Positive
ini mksdnya gmna ya bang	Neutral
bang apa kalo nyimpan uang di jenius dapat terpotong tiap bulannya atau gimna	Neutral
buat maxi safe gimana caranya	Positive
tanya saya soal jenius apk kampret duit saya juta gak bisa di ambil di bloki	Negative

#### 4.4 Ekstraksi Sentimen

Pada tahapan ini dapat dilihat gambar 6. Menunjukkan distribusi sentimen analisis dari dataset sebanyak 856 komentar dengan sebaran sentiment *neutral* 93.8%, *positive* 50.1 % negative 1.1 %. Berikut gambar sebaran pada tampilan diagram *pie*.



Gambar 6. Persentase Kemunculan Sentimen

#### 4.5 Wordcloud

Studi ini menemukan topik atau kata yang paling sering muncul dalam dataset komentar YouTube. Gambar 6 menunjukkan frekuensi kata terbanyak.



Gambar 6. Hasil Visualisasi Wordcloud

#### 4.6 Klasifikasi

Berikut ini pada gambar 7. menunjukkan hasil dari *Precision*, dan *Recall*, *F1-Score*, AUC dari klasifikasi. Laporan klasifikasi ini memperlihatkan kinerja model dalam mengklasifikasikan sentimen menjadi negatif, netral, dan positif. Model menunjukkan performa yang sangat baik dalam mendeteksi sentimen netral, dengan precision 0.94, recall 1.00, dan f1-score 0.97. Namun, model tidak berhasil mendeteksi sentimen negatif dan positif, yang terlihat dari *precision*, *recall*, dan *f1-score* yang semuanya nol untuk kedua kelas tersebut. Akurasi keseluruhan model adalah 0.94, tetapi ini terutama karena kelas netral mendominasi data. Rata-rata tertimbang menunjukkan *precision* 0.88, *recall* 0.94, dan *f1-score* 0.90, mengindikasikan kinerja keseluruhan yang baik namun tidak seimbang.

```
Classification Report:
              precision    recall  f1-score   support

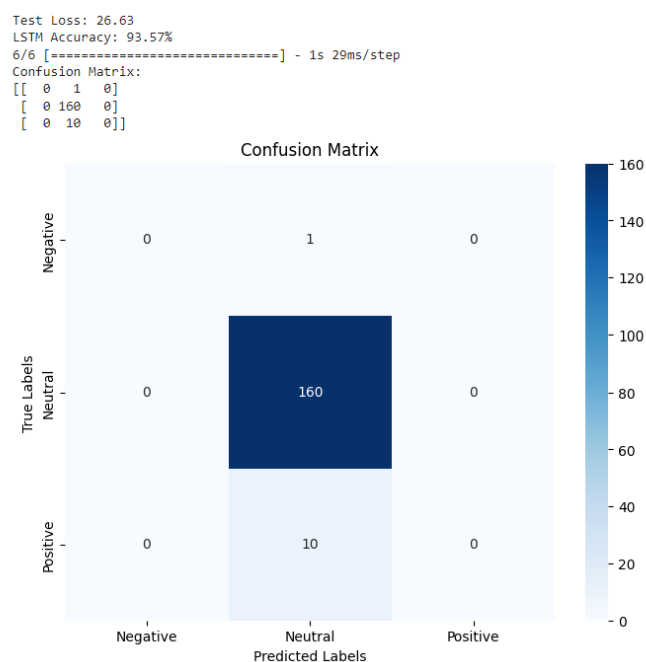
   Negative      0.00      0.00      0.00         1
   Neutral      0.94      1.00      0.97        160
   Positive      0.00      0.00      0.00         10

 accuracy              0.94        171
 macro avg              0.31        171
 weighted avg           0.88        171
```

Gambar 7. Hasil Klasifikasi

#### 4.7 Evaluasi

Dalam proses pelatihan ini, hasil *Confusion Matrix* untuk semua kelas, yaitu positif, negatif, dan netral, dapat dilihat pada gambar 8. yang menunjukkan matriks kebingungan (*confusion matrix*) dari model LSTM yang digunakan untuk klasifikasi sentimen. Model ini memiliki akurasi sebesar 93.57% dengan nilai kehilangan (*loss*) sebesar 26.63. Dari matriks kebingungan, terlihat bahwa model sangat baik dalam mengklasifikasikan sentimen netral dengan 160 dari 160 prediksi benar. Namun, model gagal mengklasifikasikan sentimen negatif dan positif dengan benar, karena semua sampel negatif dan positif diklasifikasikan sebagai netral. Ini menunjukkan bahwa model memiliki ketidakseimbangan dalam mendeteksi kelas sentimen yang jarang muncul (negatif dan positif) dibandingkan dengan kelas sentimen netral yang dominan.



Gambar 8. Hasil *Confusion Matrix*

## 5. Simpulan

Dalam penelitian, model LSTM yang dikembangkan menunjukkan tingkat akurasi yang impresif sebesar 93.57% pada set data uji, menonjol dalam mengenali kategori netral dengan *precision* dan *recall* masing-masing sebesar 0.94. Namun, model tersebut mengalami kesulitan signifikan dalam mengidentifikasi kategori negatif dan positif, dengan kedua kategori tersebut memiliki *precision* dan *recall* sebesar 0. Hal ini mengindikasikan kegagalan model dalam menggeneralisasi pembelajaran terhadap kelas-kelas yang kurang representasi dalam data Latihan.

Ketidakeimbangan dalam distribusi kelas tampaknya menjadi faktor krusial yang mempengaruhi kinerja model, dimana model cenderung bias terhadap kelas dengan jumlah sampel terbanyak. Hasil ini menggarisbawahi pentingnya mengimplementasikan strategi penanganan ketidakseimbangan kelas, seperti teknik resampling atau penerapan pembobotan kelas selama pelatihan, untuk meningkatkan sensitivitas model terhadap kelas minoritas. Kesimpulan ini menegaskan perlunya pendekatan yang lebih holistik dalam preprocessing data dan pengembangan model yang tidak hanya fokus pada optimasi akurasi secara keseluruhan tetapi juga pada kemampuan model dalam mengklasifikasikan semua kategori secara adil dan seimbang. Kesimpulan ini juga memicu pertimbangan untuk studi lebih lanjut yang mengarah pada penyesuaian model dan pengujian dengan data yang lebih beragam untuk memvalidasi temuan ini.

## Referensi

- [1] J. Shobana & M. Murali, "Complex & Intelligent Systems," *An Effic. Sentim. Anal. Methodol. based long short-term Mem. networks*, vol. Volume 7, p. pages 2485–2501, 2021, doi: <https://doi.org/10.1007/s40747-021-00436-4>.
- [2] M. A. Nurrohmat and A. SN, "Sentiment Analysis of Novel Review Using *Long Short-Term Memory* Method," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 13, no. 3, p. 209, 2019, doi: 10.22146/ijccs.41236.
- [3] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [4] K. Gopalakrishnan and F. M. Salem, "Sentiment Analysis Using Simplified *Long Short-Term Memory* Recurrent Neural Networks," arXiv, pp. 1–6, 2020.
- [5] A.K. Sinha, "Bank Complaint Analysis Using Deep Learning LSTM." [Online]. Available: <https://github.com/Akashsinha1999/Bank-Complaint-Analysis-using-LSTM-Deep-Learning-Algorithm>
- [6] A.M. Ozbayoglu, M.U. Gudelek, & O.B. Sezer, "Deep learning for financial applications: A survey. *Applied soft computing*, 93, 106384,2020.
- [7] X. Li, H. Xie, L. Chen, J. Wang, and X. Deng, "Knowledge-Based Systems News impact on stock price return via sentiment analysis," vol. 69, pp. 14–23, 2014.
- [8] A. Ray and A. Chakrabat, "A combined approach of the Rule-based and Deep Learning method for aspect level sentiment analysis." [Online]. Available: <https://link.springer.com/content/pdf/10.1007/s40747-021-00436-4.pdf>
- [9] S.I. Muhamad, "Sentimen Analisis Pada Media Sosial Menggunakan Metode Naive Bayes Classifier (Nbc)," *Teknologipintar.org*, vol. 3, no. 2, pp. 1-9, 2023.
- [10] G. Y. Lee, L. Alzamil, B. Doskenov, and A. Termehchy, "Model Performance," pp. 1–6, 2021.
- [11] C. A. Misrun, F. Sains, D. A. N. Teknologi, U. Islam, N. Sultan, and S. Kasim, "Terhadap Anies Baswedan Sebagai Bakal Calon Presiden 2024 Menggunakan Metode Naive Bayes Classifier," Repository, uin-suska, 2024.
- [12] D. Jurafsky and J. H. Martin, "Summary of Contents," New York: Cambridge university press, 2023.
- [13] R. Slamet, W. Gata, A. Novtariany, K. Hilyati, and F. A. Jariyah, "Analisis Sentimen Twitter Terhadap Penggunaan Artis Korea Selatan Sebagai Brand Ambassador Produk Kecantikan Lokal," *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 5, no. 1, pp. 145–153, 2022, doi: 10.31539/intecomsv5i1.3933.
- [14] H. H. S. Gunawan, I. N. Saniyah, "Improving the accuracy of text classification using *Stemming* method, a case of non-formal Indonesian conversation," *J. Big Data*, vol. 8, no. 1, pp. 1–20, 2021.
- [15] S. Chanda & P. Sukomal "The Effect of Stopword Removal on Information Retrieval for

- Code-Mixed Data Obtained Via Social Media,” *SN Comput. Sci.*, vol. Volume 4, no. 3, pp. 113-124, 2023, doi: <https://doi.org/10.1007/s42979-023-01942-7>.
- [16] T.B. Sianturi, I. Cholissodin, & N. Yudistira, "Penerapan Algoritma Long Short-Term Memory (LSTM) berbasis Multi Fungsi Aktivasi Terbobot dalam Prediksi Harga Ethereum. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vo;. 7, no. 3, pp. 1101-1107, 2023.
- [17] A. E. Augustia, R. Taufan, Y. Alkhalifi, and W. Gata, "Analisis Sentimen Omnibus Law Pada Twitter Dengan Algoritma Klasifikasi Berbasis Particle Swarm Optimization," vol. 23, no. 2, pp. 59-68, 2021.
- [18] K. S. Nugroho *et al.*, "Deteksi Depresi Dan Kecemasan Pengguna Twitter," *Ciastech*, pp. 287–296, 2021.