

**ANALISIS SENTIMEN PADA PENGGUNA BANK JAGO DI
SOSIAL MEDIA TWITTER DENGAN MENGGUNAKAN
ALGORITMA BERT**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana Program
Studi Informatika



disusun oleh

HAFEZ AR RAASYID SOEPRIYONO

19.11.3289

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

**ANALISIS SENTIMEN PADA PENGGUNA BANK JAGO DI
SOSIAL MEDIA TWITTER DENGAN MENGGUNAKAN
ALGORITMA BERT**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana Program
Studi Informatika



disusun oleh

HAFEZ AR RAASYID SOEPRIYONO

19.11.3289

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

SKRIPSI

ANALISIS SENTIMEN PADA PENGGUNA BANK JAGO DI SOSIAL MEDIA TWITTER DENGAN MENGGUNAKAN ALGORITMA BERT

yang disusun dan diajukan oleh

Hafez Ar Raasyid Soepriyono

19.11.3289

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 Mei 2025

Dosen Pembimbing,



Theopilus Bayu Sasongko, S.Kom, M.Eng
NIK. 190302375

HALAMAN PENGESAHAN

SKRIPSI

ANALISIS SENTIMEN PADA PENGGUNA BANK JAGO DI SOSIAL MEDIA TWITTER DENGAN MENGGUNAKAN ALGORITMA BERT

yang disusun dan diajukan oleh

Hafez Ar Raasyid Soepriyono

19.11.3289

Telah dipertahankan di depan Dewan Penguji
pada tanggal 23 Mei 2025

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Theopilus Bayu Sasongko, S.Kom., M.Eng.
NIK. 190302375

Rumini, S.Kom., M.Kom.
NIK. 190302246

Dina Maulina, S.Kom., M.Kom.
NIK. 190302250



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 23 Mei 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusrini, M.Kom.
NIK. 1903020106

PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Hafez Ar Raasyid Soepriyono
NIM : 19.11.3289

Menyatakan bahwa Skripsi dengan judul berikut:

Analisis Sentimen Pada Pengguna Bank Jago di Sosial Media Twitter Dengan Menggunakan Algoritma BERT

Dosen Pembimbing: Theopilus Bayu Sasongko, S.Kom, M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 23 Mei 2025

Yang Menyatakan,



Hafez Ar Raasyid Soepriyono

HALAMAN PERSEMBAHAN

Alhamdulillahirrobbal `alamin, skripsi ini penulis persembahkan untuk:

1. Allah SWT, karena berkat rahmad dan hidayah-Mu, saya masih diberikan kesempatan untuk dapat memperjuangkan apa yang perlu diperjuangkan.
2. Papa, Mama, dan Kakak-kakak tercinta yang telah memberikan motivasi dan dukungan material serta moral yang tiada tara.
3. Bapak Rektor Universitas Amikom Yogyakarta dan staf ahli, sehingga saya dapat menimba ilmu di universitas ini.
4. Dosen pembimbing, Bapak Theopilus Bayu Sasongko, S.Kom, M.Eng., yang telah membimbing dari awal hingga terselesaikannya skripsi, serta seluruh dosen pengajar Universitas Amikom Yogyakarta khususnya Jurusan Informatika yang telah memberikan ilmu yang bermanfaat hingga pada tahap ini.
5. Semua teman kelas 19-IF-11 terutama kepada Faisal, Ibrahim, Alfian yang telah banyak memberi banyak bantuan, dukungan, dan motivasi selama perkuliahan.
6. Teman-teman pengurus dan kader HMIF periode 2022/2023 yang sudah memberikan banyak dukungan dan pengalaman yang bermanfaat, serta kebersamai selama menjadi mahasiswa.
7. Para sahabat Teletubbies Rizput, Zulkifli, dan Desthina yang telah memberi dukungan dan selalu menemani dalam suka maupun duka.
8. Kepada Mardiah, Kusuma, Mas David, dan Mas Danu yang telah memberikan dukukangan dan dorongan untuk menyelesaikan skripsi.
9. Para teman-teman virtual dan komunitas di sosial media twitter terutama pada *Base @collegemfs* yang telah memeberikan inspirasi, motivasi, dan dukungan pada penulis
10. Semua pihak yang turut membantu saya secara langsung maupun tidak langsung yang tidak dapat saya sebutkan satu persatu, terima kasih.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT, karena dengan Rahmat dan hidayah-Nya, penulis dapat menyelesaikan penyusunan skripsi dengan judul “Analisis Sentimen Pada Pengguna Bank Jago di Sosial Media Twitter Dengan Menggunakan Algoritma BERT”. Skripsi ini disusun dalam rangka pemenuhan persyaratan akademis guna meraih gelar Sarjana Komputer pada Universitas Amikom Yogyakarta.

Penulis ingin menyampaikan ucapan terimakasih yang sebesar-besarnya kepada dosen pembimbing, Bapak Theopilus Bayu Sasongko, S.Kom, M.Eng dan [Dosen Penguji], atas waktu dan masukan yang sangat berharga dalam pengembangan skripsi ini.

Penulis juga ucapkan terimakasih kepada semua pihak yang turut serta mendukung dan memberikan semangat. Ucapan terimakasih yang tak terhingga juga disampaikan kepada orang tua, keluarga, dan pihak-pihak yang telah memberikan dukungan moral dan doa restu selama penyusunan skripsi ini.

Semoga hasil dari skripsi ini dapat memberikan manfaat dan menjadi kontribusi positif dalam pengembangan ilmu pengetahuan dan teknologi. Akhir kata, penulis menyadari bahwa skripsi ini masih jauh dari kata kesempurnaan. Oleh karena itu, kritik dan sarah yang bersifat membangun sangat diharapkan demi perbaikan pada penelitian mendatang.

Yogyakarta, 8 April 2025

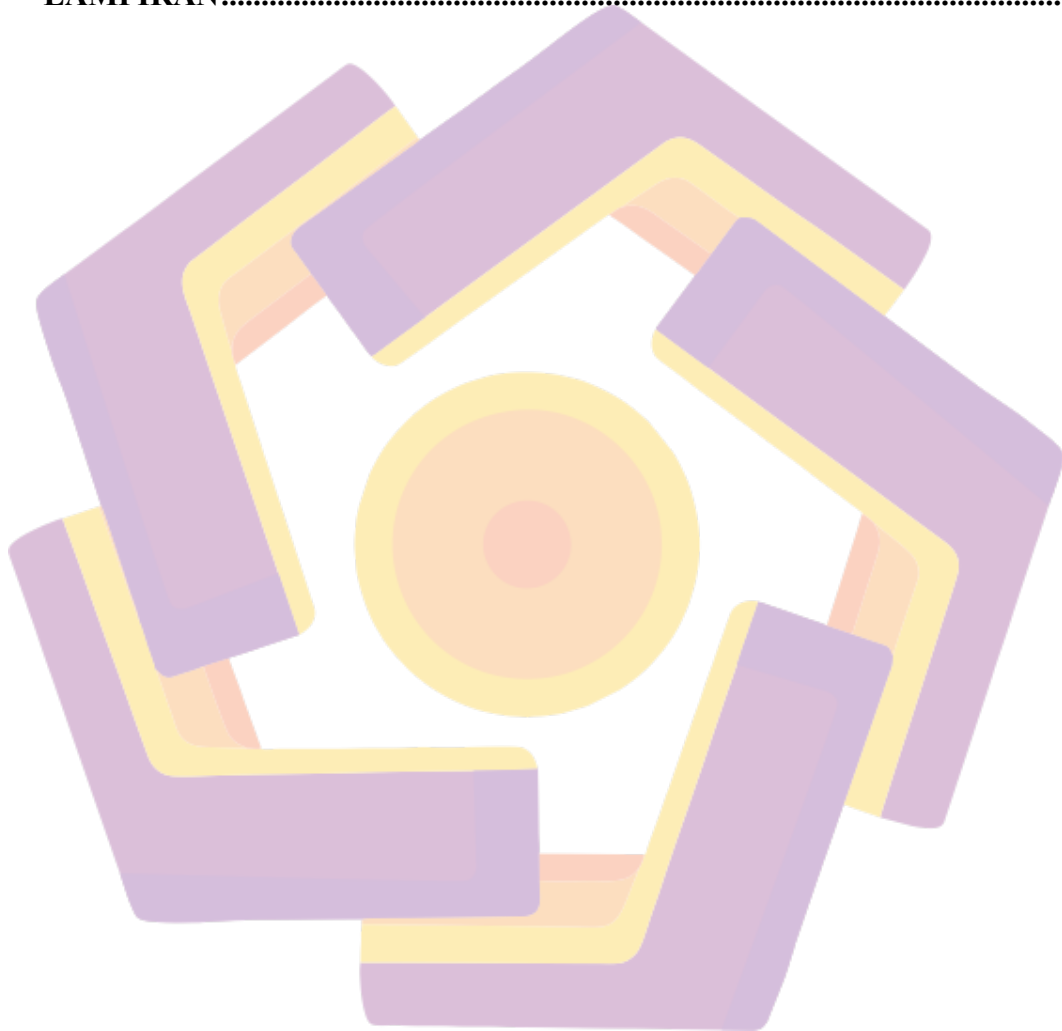
Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
PERNYATAAN KEASLIAN SKRIPSI.....	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR.....	x
DAFTAR TABEL	xii
INTISARI	xiii
ABSTRACT	xiv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah.....	4
1.4. Tujuan Penelitian.....	4
1.5. Manfaat Penelitian.....	4
1.6. Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI	6
2.1. Tinjauan Pustaka	6
2.2. Dasar Teori	14
2.2.1. Bank Jago	14
2.2.2. Text Mining.....	14
2.2.3. Analisis Sentimen	15
2.2.4. <i>Natural Language Processing</i> (NLP)	17
2.2.5. Deep Learning.....	18
2.2.6. <i>Bidirectional Encoder Reprerentation Form Transformers</i> (BERT). 20	
2.2.7. Confusion Matrix	29

2.2.8.	Python	31
2.2.9.	Flask	31
BAB III METODE PENELITIAN		33
3.1.	Objek Penelitian	33
3.2.	Alur Penelitian.....	33
3.3.1	Pengumpulan Data	34
3.3.2	Pelabelan Data.....	34
3.3.3	<i>Pre-Processing</i> Data	35
3.3.4	EDA	38
3.3.5	Data Preparation.....	38
3.3.6	Modeling	39
3.3.7	Prediksi dan Evaluasi	40
3.3.	Alat dan Bahan Penelitian	40
3.3.1.	Alat.....	40
3.3.2.	Bahan.....	40
BAB IV HASIL DAN PEMBAHASAN		41
4.1	Hasil Pengumpulan Data	41
4.2	Pelabelan Data	42
4.3	Pre-processing Data.....	43
4.4	EDA.....	47
4.4.1	Persentase.....	47
4.4.2	Visualisasi (Word Cloud).....	48
4.5	Data <i>Preparation</i>	50
4.5.1	Tokenizing and Encode.....	51
4.5.2	Split Data.....	51
4.5.3	Data <i>Loader</i>	52
4.6	<i>Modeling</i>	52
4.6.1	<i>Set-up BERT Pretrained Model</i>	52
4.6.2	<i>Setting-up Optimizer and Scheduler</i>	53
4.6.3	<i>Training BERT</i>	53
4.7	Evaluasi dan Prediksi	56

4.8 Hasil <i>Deploy</i>	62
BAB V PENUTUP	64
5.1. Kesimpulan	64
5.2. Saran	64
DAFTAR PUSTAKA	66
LAMPIRAN	70



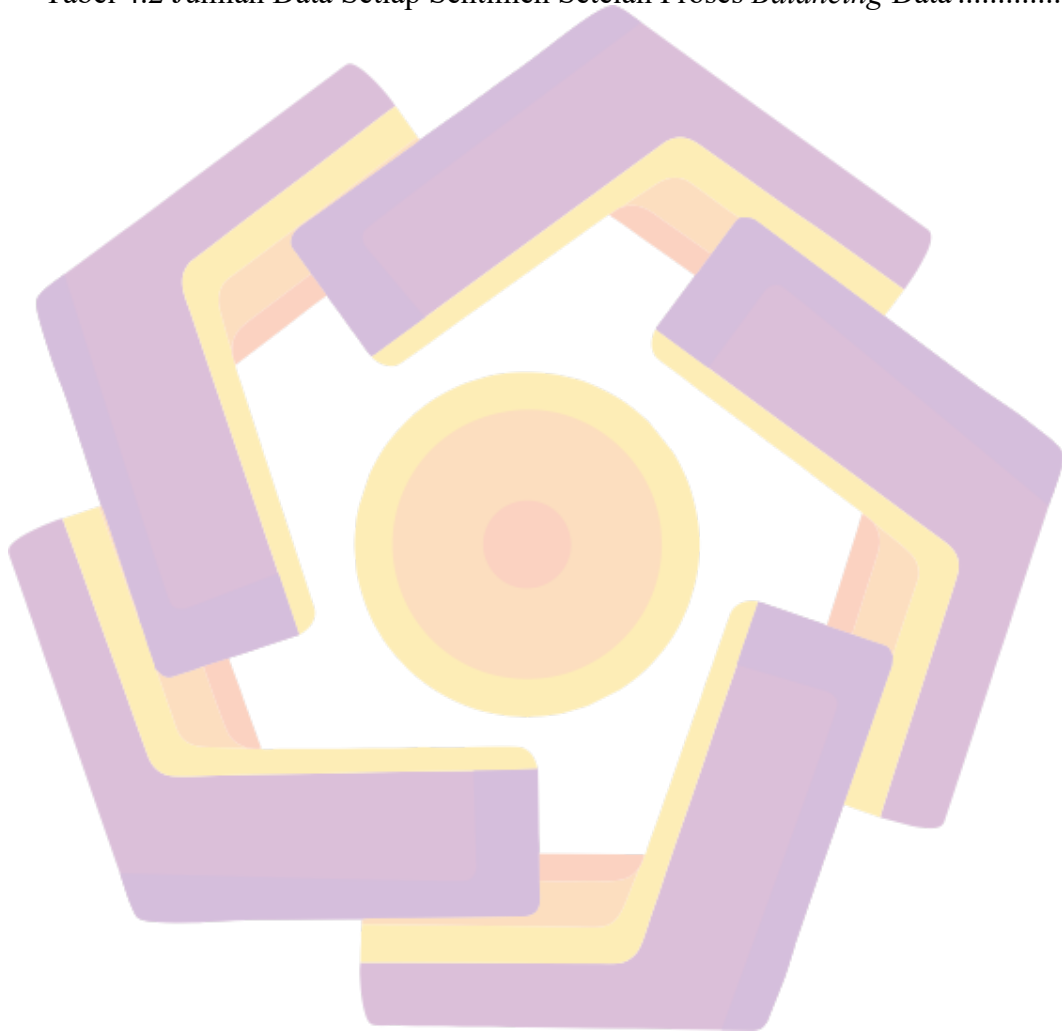
DAFTAR GAMBAR

Gambar 2.1 Jaringan Saraf <i>Feed-forward</i>	19
Gambar 2.2 Encoder dan decoder	22
Gambar 2.3 Proses Self-attention.....	23
Gambar 2.4 Proses Encoder	23
Gambar 2.5 Variasi Ukuran pada BERT.....	24
Gambar 2.6 Arsitektur BERT	25
Gambar 2.7 Perbedaan antara arsitektur BERT dengan OpenAI GPT dan ELMo.....	26
Gambar 2.8 Proses pada MLM	27
Gambar 2.9 Proses Pre-training pada BERT.....	27
Gambar 2.10 Representasi input pada BERT	28
Gambar 3.1 Diagram Alur Penelitian.....	33
Gambar 3.2 Alur <i>Scraping</i> Data Tweet	34
Gambar 3.3 Alur Labeling Data.....	35
Gambar 3.4 Alur <i>Case Folding</i>	36
Gambar 3.5 Alur <i>Filtering</i>	36
Gambar 3.6 Alur <i>Tokenizing</i>	37
Gambar 3.7 Alur <i>Stemming</i>	37
Gambar 4.1 Tweet Mengenai Ulasan Pengguna Bank Jago	41
Gambar 4.2 Contoh Hasil Pengumpulan Data	42
Gambar 4.3 Hasil Pelabelan Data	43
Gambar 4.4 Hasil Case Folding Data.....	43
Gambar 4.5 Hasil <i>Filtering</i> Data.....	44
Gambar 4.6 Hasil Tokenizing Data.....	45
Gambar 4.7 Hasil <i>Stemming</i> Data	45
Gambar 4.8 Hasil Klasifikasi Data.....	47
Gambar 4.9 Hasil Persentase Data	48
Gambar 4.10 Visualisasi Kategori Positif.....	49
Gambar 4.11 Visualisasi Kategori Netral	49

Gambar 4.12 Visualisasi Kategori Negatif	50
Gambar 4.13 Tokenisasi dengan Token Khusus.....	51
Gambar 4.14 Hasil Menambahkan <i>Padding</i>	51
Gambar 4.15 Hasil <i>Set-up Pretrained Model</i>	53
Gambar 4.16 Grafik <i>Training Loss</i> Skenario Pertama.....	54
Gambar 4.17 Grafik <i>Training Loss</i> Skenario Kedua	54
Gambar 4.18 Grafik <i>Training Loss</i> Skenario Ketiga	55
Gambar 4.19 Diagram <i>Confusion Matrix</i> Skenario Pertama	56
Gambar 4.20 Akurasi Analisis Ulasan Skenario Pertama.....	57
Gambar 4.21 Diagram <i>Confusion Matrix</i> Skenario Kedua.....	58
Gambar 4.22 Akurasi Analisis Ulasan Skenario Kedua	59
Gambar 4.23 Diagram <i>Confusion Matrix</i> Skenario Ketiga.....	60
Gambar 4.24 Akurasi Analisis Ulasan Skenario Ketiga	60
Gambar 4.25 Tampilan Sistem Menampilkan Sentimen Negatif	62
Gambar 4.26 Tampilan Sistem Menampilkan Sentimen Netral	63
Gambar 4.27 Tampilan Sistem Menampilkan Sentimen Positif.....	63

DAFTAR TABEL

Tabel 2.1 Perbandingan Tinjauan Pustaka	8
Tabel 2.1 <i>Confusion Matrix</i>	29
Tabel 4.1 Jumlah Data Setiap Sentimen	46
Tabel 4.2 Jumlah Data Setiap Sentimen Setelah Proses <i>Balancing</i> Data	47



INTISARI

Beberapa bank sentral, termasuk Bank Indonesia (BI) untuk saat ini sedang dalam proses mengembangkan rupiah digital atau yang biasa dikenal sebagai *Central Bank Digital Currency* (CBDC). CBDC merupakan bentuk dari uang digital yang dibuat dan diatur oleh bank sentral. Bank Jago merupakan salah satu bank digital yang ada di Indonesia dan merupakan anak perusahaan dari aplikasi gojek. Dari banyaknya nasabah atau pengguna yang dimiliki oleh bank jago ini mendapat berbagai macam respon yang di terima dari berbagai kalangan baik itu respon positif, negatif maupun netral dan sebagian besar pengguna lebih nyaman untuk menyampaikan respon tersebut melalui sosila media terutama pada aplikasi Twitter. Maka dari itu di perlukan sebuah analisa sentimen untuk dapat menentukan respon-respon tersebut.

Dalam menentukan apakah opini dan sentimen tersebut memiliki respon positif, negatif maupun netral maka diperlukan sebuah klasifikasi. Klasifikasi ini merupakan tahap utama proses dalam analisis sentimen untuk memahami pola sentimen dari data yang telah diproses. Sebelum dilakukan klasifikasi maka akan dilakukan pengambilan data terlebih dahulu dari tweet atau postingan pada media sosial twitter dengan menggunakan kata kunci Bank Jago. Setelah itu data yang didapat kemudian diolah sehingga dapat menentukan klasifikasi opini dan sentimen dari postingan para nasabah tersebut apakah bernilai positif, negatif, dan netral yang pada kali ini klasifikasi akan menggunakan metode BERT (Bidirectional Encoder Representation from Transformers).

Berdasarkan hasil analisis data sebelum diterapkan teknik penyeimbangan dengan SMOTE, distribusi sentimen pada ulasan pengguna Bank Jago menunjukkan ketidakseimbangan yang cukup signifikan. Sentimen netral mendominasi dengan jumlah 1177 data, sedangkan sentimen negatif dan positif jauh lebih sedikit, masing-masing berjumlah 578 dan 351 data. Ketidakseimbangan ini dapat berdampak pada performa model analisis sentimen, di mana model cenderung lebih akurat dalam mengenali sentimen netral dibandingkan dengan sentimen positif atau negatif. Oleh karena itu, diperlukan metode seperti SMOTE untuk menyeimbangkan jumlah data di setiap kategori sentimen guna meningkatkan akurasi dan keandalan model dalam mengklasifikasikan sentimen pengguna secara lebih adil. Penelitian ini menerapkan model BERT untuk analisis sentimen ini menggunakan tiga skenario pembagian data dengan tingkat akurasi terbaik mencapai 85%.

Kata Kunci: Analisis sentimen, Bank digital, Bank Jago, BERT

ABSTRACT

Several central banks, including Bank Indonesia (BI) are currently in the process of developing a digital rupiah or what is commonly known as the Central Bank Digital Currency (CBDC). CBDC is a form of digital money created and regulated by the central bank. Bank Jago is one of the digital banks in Indonesia and is a subsidiary of the Gojek application. From the many customers or users owned by Bank Jago, they have received various responses from various groups, both positive, negative and neutral responses and most users are more comfortable conveying these responses through social media, especially on the Twitter application. Therefore, a sentiment analysis is needed to determine these responses.

In determining whether the opinion and sentiment have a positive, negative or neutral response, a classification is needed. This classification is the main stage of the sentiment analysis process to understand the sentiment pattern of the data that has been processed. Before the classification is carried out, data will be taken first from tweets or posts on Twitter social media using the keyword Bank Jago. After that, the data obtained is then processed so that it can determine the classification of opinions and sentiments from the customer's posts, whether they are positive, negative, and neutral, which this time the classification will use the BERT (Bidirectional Encoder Representation from Transformers) method.

Based on the results of data analysis before applying the balancing technique with SMOTE, the distribution of sentiment in Bank Jago user reviews shows a significant imbalance. Neutral sentiment dominates with a total of 1177 data, while negative and positive sentiments are much less, totaling 578 and 351 data respectively. This imbalance can have an impact on the performance of the sentiment analysis model, where the model tends to be more accurate in recognizing neutral sentiments compared to positive or negative sentiments. Therefore, a method such as SMOTE is needed to balance the amount of data in each sentiment category in order to increase the accuracy and reliability of the model in classifying user sentiments more fairly. The study implementing the BERT model for sentiment analysis used three data splitting scenarios, with the highest accuracy reaching 85%.

Keywords: *Sentiment analysis, Digital bank, Bank Jago, BERT*