

Penerapan Model LSTM pada Analisis Sentimen Ulasan Pengguna Aplikasi *Shopee Google Play Store*

Rahayu Noveandini¹, Maria Sri Wulandari², Farhan Rasyad³

^{1,2} Sistem Informasi, Sistem Informasi, STMIK Jakarta STI&K

³ Teknik Informatika, Universitas Gunadarma

ayu.noveandini@gmail.com*, mwuland@yahoo.com, farsyad13@gmail.com

Abstract

In recent years, technology-based marketing has experienced a significant growth spurt. This is in line with the rapid advancements in digital technologies, such as artificial intelligence (AI), big data, machine learning, and the Internet of Things (IoT). These technologies allow companies to recognize consumer behavior more accurately, personalize marketing messages, and optimize sales strategies in real time. The emergence of digital platforms such as social media, e-commerce has also accelerated this growth. Companies no longer rely solely on traditional marketing methods but instead build marketplaces that are digital markets. To measure the level of customer satisfaction, improve the quality of services and products, detect problems and support data-based business decisions, an understanding, opinion or perception of user satisfaction with the services or products offered on the platform is needed. Sentiment analysis aims to understand a person's attitude or feelings towards something through what is contained in a text whether it is positive, negative or neutral. The LSTM (Long ShortTerm Memory) algorithm is one of the models for performing sentiment analysis. In this study, the LSTM algorithm is used with the combination of the SMOTE and BERT methods to measure the level of accuracy, precision and recall of user reviews of the Shopee Google Playstore application. This research is divided into 3 batches, namely batches 32, 64 and 128 with an epoch value of 5 and is able to provide an accuracy rate performance of 99.56% with a negative sentiment precision value of 1.00, positive sentiment of 1.00 and a recall value of 0.99 in batch 128.

Keywords: marketing, marketplace, shopee application, sentiment, LSTM, accuracy

Abstrak

Dalam beberapa tahun terakhir, pemasaran berbasis teknologi telah mengalami lonjakan pertumbuhan yang signifikan. Hal ini terjadi seiring dengan kemajuan pesat dalam bidang teknologi digital, seperti kecerdasan buatan (AI), *big data*, *machine learning*, dan *Internet of Things* (IoT). Teknologi ini memungkinkan perusahaan untuk mengenali perilaku konsumen secara lebih akurat, mempersonalisasi pesan pemasaran, serta mengoptimalkan strategi penjualan secara *real time*. Kemunculan *platform digital* seperti media sosial, *e-commerce* turut mempercepat pertumbuhan ini. Perusahaan tidak lagi hanya mengandalkan metode pemasaran tradisional melainkan dengan membangun *marketplace* yang merupakan pasar *digital*. Untuk mengukur tingkat kepuasan pelanggan, meningkatkan kualitas layanan dan produk, mendeteksi masalah serta mendukung keputusan bisnis berbasis data diperlukan suatu pemahaman, opini maupun persepsi kepuasan pengguna terhadap layanan atau produk yang ditawarkan di *platform* tersebut. Analisis sentimen bertujuan untuk mengklasifikasi sikap atau perasaan seseorang terhadap suatu hal berdasarkan suatu teks apakah bersifat positif, negatif atau netral. Algoritma LSTM (*Long ShortTerm Memory*) merupakan salah satu model untuk melakukan analisis sentimen. Penggunaan algoritma LSTM dengan penggabungan metode SMOTE dan BERT pada penelitian ini untuk mengukur tingkat akurasi, presisi dan *recall* dari ulasan review pengguna aplikasi *Shopee Google Playstore*. Pada penelitian ini dibagi menjadi 3 batch yaitu batch 32, 64 dan 128 dengan nilai *epoch* sebesar 5 dan mampu memberikan performa tingkat akurasi sebesar 99.56% dengan nilai *precision* sentiment negatif 1.00, sentiment positif 1.00 dan nilai *recall* sebesar 0.99 pada batch 128.

Kata kunci : pemasaran, marketplace, aplikasi shopee, sentimen, LSTM, akurasi

©This work is licensed under a Creative Commons Attribution -ShareAlike 4.0 International License

1. Pendahuluan

Konsep pemasaran saat ini mengalami pergeseran dalam berinteraksi dengan konsumen dari model pemasaran tatap muka menjadi pemasaran menggunakan teknologi. Hal ini disebabkan dengan perubahan teknologi digital yang semakin maju [1]. Proses bisnis dengan perubahan model tersebut mendorong transformasi ke dalam sebuah ruang pasar virtual yang dinamakan *marketplace* [2]. Hampir seluruh masyarakat Indonesia senang berbelanja menggunakan *platform online* atau yang dikenal dengan *marketplace*. *Marketplace* adalah sebuah

platform atau tempat berupa digital yang mempertemukan penjual dan pembeli. *Marketplace* dapat dikategorikan berdasarkan berbagai macam kriteria seperti berdasarkan jenis barang/jasa yang diperdagangkan, berdasarkan lokasi dalam hal ini lokal maupun global, berdasarkan skala yaitu *marketplace* skala besar dengan persaingan pasar yang kompetitif serta mempunyai jumlah penjual dan pembeli yang besar dan skala kecil yang memiliki jumlah penjual dan pembeli yang sedikit. Perubahan model belanja pada masyarakat ditandai dengan pertumbuhan *marketplace* yang semakin menunjang kebutuhan masyarakat [3].

[4] Pada *marketplace* juga disediakan berbagai macam fitur untuk membantu mempermudah transaksi seperti fitur pencarian, pembayaran, pengiriman bahkan jika penjual tidak mengirimkan barangnya hingga batas waktu yang ditentukan, maka dana yang telah dikeluarkan pembeli dapat kembali ke akun pembeli. Disamping itu juga terdapat fitur untuk melakukan tanya jawab dengan penjual, tanya jawab dengan admin platform *marketplace* serta fitur *review* pelanggan untuk melihat seberapa jauh kepuasan pelanggan terhadap produk yang ditawarkan oleh penjual. Disini terlihat bahwa *marketplace* benar – benar sebuah wadah yang cukup terjamin keamanannya dalam bertransaksi. Fitur – fitur yang disediakan oleh marketplace berfungsi untuk memudahkan transaksi penjualan, memperluas jaringan pemasaran, meningkatkan efisiensi dalam proses jual beli hingga meningkatkan persaingan antar penjual. Dengan adanya fitur *review* pelanggan maka terjadi persaingan yang cukup ketat dalam proses penjualan tersebut. Sentimen dari kepuasan pelanggan menandakan reaksi emosional atau opini pelanggan mencerminkan tingkat kepuasan terhadap suatu produk sehingga menjadi indikator penting dalam mengukur performa produk selama pelanggan menggunakan produk tersebut [5]. Analisis sentimen dimaksudkan untuk menggali wawasan dari sekumpulan ulasan yang membantu penjual untuk dapat merespon kebutuhan pelanggan dengan lebih baik [6].

Recurrent Neural Network (RNN)

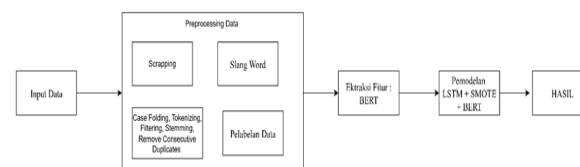
RNN (*Recurrent Neural Network*) merupakan arsitektur jaringan saraf tiruan yang bekerja secara berurutan seperti data teks, audio, data waktu (*time series*). Berbeda dengan jaringan saraf biasa yang hanya memproses input satu persatu dan tidak mengingat proses sebelumnya, RNN dapat mengingat konteks atau keadaan sebelumnya pada saat memproses data *sequential* [7]. Varian dan evolusi RNN salah satunya adalah LSTM (*Long Short Term Memory*). LSTM di rancang untuk mengatasi masalah *vanishing gradient* yaitu memiliki sel memori yang dapat menyimpan informasi dalam waktu lama dan menggunakan gerbang (*gate*) yaitu input gate, forget gate dan output gate. Struktur inti LSTM yaitu : (1) Cell State (C_t) adalah memori jangka panjang yang berfungsi aliran informasi, (2) Hidden Output (h_t) yang merupakan output yang digunakan untuk proses input selanjutnya, Gate (gerbang) yaitu forget gate (f_t), input gate (i_t) dan output gate (o_t).

Algoritma LSTM (*Long Short Term Memory*) digunakan pada penelitian analisis sentimen ini untuk memproses dan memprediksi data. Algoritma LSTM (*Long Short Term Memory*) dalam RNN diterapkan guna mengatasi kelemahan RNN dalam mengingat informasi jangka panjang dengan meningkatkan kemampuan memori, menyaring informasi yang relevan dan memberikan mekanisme ketertarikan sehingga setiap kata lebih kontekstual. Pengujian dilakukan dengan membagi data berdasarkan

penerapana teknik *stemming*, dan *non-stemming*. Akurasi sebesar 81%, F-Measure 54,27% dan nilai representatif sebesar 41% diperoleh dari hasil analisis ulasan berita *non-stemming*. Setelah dilakukan analisis *review* berita dengan metode *stemming*, sistem menghasilkan model yang memiliki tingkat akurasi sebesar 88%, F-Measure 58.20% tetapi representasi atau relevansi kata – kata setelah *stemming* hanya mencerminkan 44% dari makna asli [8]. Akurasi sebesar 83% dan nilai *loss* 0.138 dihasilkan pada analisis sentimen menggunakan LSTM pada ulasan *marketplace* Shopee. Kinerja arsitektur yang dibangun cukup baik dengan mengklasifikasi sentimen positif dan negatif namun, performanya kurang optimal pada kelas netral [9]. Eksperimen yang dilakukan terhadap dua jenis fitur data penggunaan listrik PLN yaitu siang dan malam. Untuk data siang, digunakan konfigurasi model dengan *sequence length* 20, hidden size 8, 3 lapisan LSTM, dan proporsi data pelatihan sebesar 70%, dihasilkan nilai RMSE sebesar 46.72. Sementara itu, pada data malam digunakan *sequence length* 30, hidden size 8, 1 lapisan LSTM, serta 30% data pelatihan, dengan RMSE sebesar 51.05 [10]. Penelitian ini bertujuan untuk melakukan penilaian pendapat atau tanggapan pengguna yang mampu memberikan masukan bagi pengembang dalam upaya peningkatan kualitas layanan aplikasi.

2. Metode Penelitian

Pada penelitian ini dimulai dengan tahap awal yaitu pengumpulan data *review* ulasan aplikasi shopee pada *google playstore* dalam Bahasa Indonesia, *preprocessing*, pelabelan data, ekstraksi fitur, pembuatan model dengan skema LSTM + SMOTE + BERT fitur dengan membagi data menjadi 70/30 dengan regulasi dan *dropout*. Model ini diklasifikasi pada data ulasan dan dimasukkan dalam ulasan positif dan ulasan negatif dari penggunaan aplikasi *marketplace* Shopee. Dengan skema yang tepat diharapkan mampu memberikan rekomendasi penggunaan aplikasi marketplace Shopee pada pengguna.



Gambar 1. Metode Penelitian

2.1. Pengumpulan Data

Penelitian ini memanfaatkan data sekunder dimana data sudah ada sebelum penelitian dilakukan dan telah digunakan oleh peneliti lainnya [11]. Data diambil dengan metode *scrapping* dari aplikasi *Shopee* dari *Google Playstore* [12]. Pengumpulan data teks ini dalam Bahasa Indonesia dengan metode *scrapping*. Data yang digunakan sebanyak 50.000 data ulasan penggunaan aplikasi *Shopee* dari *Google Playstore*

dalam Bahasa Indonesia. Ulasan tersebut mengandung ulasan positif, ulasan negatif dan ulasan netral dalam format *.csv. Data tersebut digunakan dalam pelatihan dan pengujian dengan pemanfaatan algoritma LSTM yang dibangun.

2.2. Data Preprocessing

Setelah melakukan pengumpulan data, proses berikutnya adalah *preprocessing* data. Tahap ini dilakukan guna mempersiapkan data agar dapat digunakan pada proses selanjutnya. Aktivitas yang dilakukan pada tahap ini meliputi : *Delete Null Value*: tahap ini melakukan pengecekan dan menghapus kolom yang memiliki nilai null sehingga dapat menyebabkan error pada saat proses pelatihan dan pengujian model, *Word Tokenizing*: proses memecah kalimat menjadi elemen yang lebih kecil yang biasa disebut dengan frasa, *Case Folding*: merupakan langkah normalisasi teks yang menyederhanakan variasi penulisan menjadi seragam untuk proses selanjutnya, *Stopword Removal*: merupakan proses penghapusan kata umum yang sering muncul, dan tidak memiliki kontribusi yang berarti, *Stemming*: merupakan proses pengembalian kata – kata ke bentuk dasar dengan menghilangkan imbuhan, *Slang Word*: pada tahap ini membuat kamus kata slang dan mengganti kata slang menggunakan kamus slang, *Remove Consecutive Duplicates*: tahap ini dilakukan penghapusan kata duplikat yang muncul berurutan, *Split Data*: Merupakan proses pembagian data dengan 70% data pelatihan dan 30% untuk data pengujian berdasarkan total jumlah data. Tahap ini dikerjakan guna melatih model agar dapat mengenali pola sehingga algoritma yang diterapkan mendapatkan akurasi yang baik.

2.3. Pelabelan Data

Fungsi pelabelan data dalam proses LSTM adalah untuk memberikan target keluaran (label) yang akan digunakan sebagai tabel acuan. Pelabelan data dilakukan untuk memberikan identifikasi informasi. Kolom ulasan merupakan bagian data yang diberi label. Penelitian ini akan memberikan label positif, negatif dan netral.

Secara rinci proses pelabelan pada LSTM adalah :

1. Menentukan *ground truth* yang akan menghasilkan (a) Label adalah "jawaban benar" yang harus dipelajari oleh model. (b) Pada *supervised learning* termasuk LSTM, label digunakan untuk menghitung *error* antara prediksi dengan nilai sebenarnya.

2. Mengarahkan proses training dengan ketentuan (a) LSTM akan menghasilkan output berdasarkan input data sebenarnya. (b) Label digunakan untuk menghitung *loss function* (misalnya MSE, *Cross Entropy*) sehingga bobot jaringan dapat diperbaharui dengan *backpropagation through time*.

3. Menentukan tipe masalah

Kegunaan label pada penentuan tipe masalah adalah (a) Pada klasifikasi label biasanya berupa kategori (0,1,2...), (b) Pada regresi / prediksi nilai label berupa nilai numerik (contoh harga saham esok hari)

4. Membantu evaluasi model. Kegunaan label adalah untuk mengukur akurasi, *error rate* atau metrik lain untuk menilai kinerja model.

Sehingga secara umum label berfungsi untuk menyediakan target yang digunakan LSTM agar dapat "belajar" memetakan pola dari data masukan ke hasil keluaran yang benar. Tanpa label LSTM hanya bisa digunakan untuk *unsupervised learning* yang sifatnya berbeda.

2.4. Distribusi Sentimen

Distribusi sentimen dilakukan setelah melakukan pelabelan data. Tujuannya adalah untuk melihat distribusi sebaran pelabelan data yang dibagi menjadi data ulasan positif, negatif dan netral. Distribusi probabilitas juga menunjukkan seberapa besar keyakinan model terhadap masing – masing label. Fungsi distribusi sentimen pada LSTM adalah : (a) Mempresentasikan ketidakpastian prediksi yaitu memberikan informasi tidak hanya pada kelas akhir tetapi juga seberapa model yakin terhadap pilihannya. (b) Dasar pengambilan keputusan berbasis probabilitas, misalnya ambang batas (*threshold*) dapat diatur untuk menentukan prediksi cukup yakin, (c) Menggambarkan nuansa teks, sebuah teks dapat mengandung unsur positif dan negatif sekaligus, (c) Sebagai bahan analisis lanjutan. Distribusi dapat digunakan untuk *ensemble learning* atau *sentiment trend analysis*.

Manfaat distribusi sentimen pada LSTM adalah (a) Analisis lebih banyak memberikan informasi dari sebuah label, (b) Memudahkan interpretasi model yaitu dapat dipakai untuk memeriksa apakah model ragu – ragu atau yakin, (c) Digunakan untuk filtering, (d) Mendukung *active learning*.

Kegunaan distribusi sentimen pada LSTM adalah sebagai (a) *Customer feedback monitoring* yaitu melihat proporsi positif /negatif untuk mengukur kepuasan pelanggan, (b) Media monitoring yaitu mengukur opini publik terhadap isu tertentu berdasarkan sebaran probabilitas sentimen, (c) E-commerce dan rekomendasi yaitu menentukan produk mana yang paling banyak mendapatkan sentimen positif / negatif, (d) *Chatbot* dan moderasi *content* yaitu menggunakan distribusi probabilitas untuk memutuskan respon atau memfilter komentar, (e) *Trend analysis* yaitu untuk mengamati perubahan distribusi sentimen seiring waktu [13][14].

2.5. Ekstraksi Fitur

Ekstraksi fitur adalah mengubah data mentah berurutan menjadi representasi vektor (feature representation) yang memuat pola dan informasi penting dari urutan tersebut. LSTM memiliki sel memori dan tiga gates, yaitu (a) *Forget gate* untuk menentukan informasi lama

mana yang dibuang, (b) *Input gate* untuk menentukan informasi baru mana yang akan dibuang dan (c) *Output gate* untuk menentukan bagian memori yang dikeluarkan untuk langkah berikutnya.

Secara umum proses ekstraksi fitur adalah :

1. Input berurutan masuk ke LSTM satu per satu (*time step*)
2. Sel memori memproses informasi dengan mempertahankan pola jangka panjang dan jangka pendek.
3. *Hidden state* pada langkah terakhir akan menjadi vektor.
4. Vektor fitur yang dihasilkan dapat (a) Digunakan untuk klasifikasi / regresi, (b) Diberikan ke lapisan lainnya (misal *dense layer*), (c) disimpan untuk analisis pola.

Ekstraksi fitur pada penelitian ini menggunakan pretrained model yaitu BERT. Kemudian selanjutnya dilakukan skema pengujian dengan LSTM + SMOTE + BERT dengan perbandingan presentase 70% data pelatihan dan 30% data test.

2.6. Evaluasi Hasil

Tahap ini bertujuan untuk menilai seberapa baik model analisis sentimen bekerja dengan mengamati hasil prediksi dan membandingkan dengan data sebenarnya untuk mengetahui tingkat akurasi, keandalan dan keseimbangan antara presisi dan *recall*. Untuk menghitung *precision* : tingkat keterpatan model dalam membuat prediksi positif yang benar untuk menghindari kesalahan terlalu banyak positif, *recall* : mengukur kemampuan model dalam mendeteksi atau menangkap seluruh data yang termasuk dalam kategori yang dicari tanpa melewatkannya dalam suatu dataset, *fi-score* : metrik evaluasi yang menggabungkan *precision* dan *recall* menjadi satu nilai, sehingga berguna untuk menjaga keseimbangan antara keduanya dalam analisis sentimen. dan *support* adalah jumlah data aktual dari setiap kelas dalam dataset (bukan prediksi) rumusnya adalah:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

dimana TP : True Positif, jumlah data yang termasuk kategori positif dan diprediksi positif.

FP : False Positif, jumlah data yang sebenarnya bukan positif dan diprediksi positif.

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

TP: True Positif, jumlah data positif yang diklasifikasi dengan benar.

FN: False Negatif, jumlah data positif yang salah diklasifikasi sebagai negatif.

$$Fi-Score = \frac{2 \times Precision \times Recall}{Precision+Recall} \quad (3)$$

Support = jumlah data aktual dari kelas tersebut

3. Hasil dan Pembahasan

Pada penelitian ini ditambahkan metode SMOTE dan BERT untuk mendapatkan parameter terbaik yang nilainya akan diterapkan pada model yang dibangun. Sebelum model LSTM diimplementasi, data dibagi terlebih dahulu dengan menggunakan perbandingan 70% data latih dan 30% data test.

3.1. Scrapping Data

Pada proses ini dilakukan pengambilan data dari aplikasi *Shopee* pada *google play store*. Jumlah total ulasan yang diambil sebesar 50.000 data ulasan. Pada tahap ini melakukan pengecekan jumlah nilai null di setiap kolom dengan sintak:

```
# Mengecek jumlah missing values di setiap kolom
print(df.isnull().sum())
```

Membuat kamus slang :

```
# Kamus slangword
slang_dictionary = {
    "@": "di", "abis": "habis", "wtb": "beli", "masi": "masih", "wts":
    "jual", "wtt": "tukar",
    "bgt": "banget", "maks": "maksimal", "pliss": "tolong", "bgttt":
    "banget", "indo": "indonesia",
    "bgtt": "banget", "ad": "ada", "rv": "redvelvet", "plis": "tolong",
    "pls": "tolong", "cr": "sumber",
    "cod": "bayar ditempat",
    "ajep-ajep": "dunia gemerlap", "ak": "saya", "akika": "aku",
    "akkoh": "aku", "akuwh": "aku",
    "alay": "norak", "alow": "halo", "ambilin": "ambilkan", "ancur":
    "hancur", "anjrit": "anjing",
    "anter": "antar", "ap2": "apa-apa", "apasih": "apa sih", "apes":
    "sial", "aps": "apa", "aq": "saya",
    "bandes": "bantuan desa", "bangedh": "banget",
    "banpol": "bantuan polisi", "banpur": "bantuan
    tempur",
    "basbang": "basi", "bcanda": "bercanda", "bdg":
    "bandung", "begajulan": "nakal", "beliin": "belikan",
    "bencong": "banci", "bentar": "sebentar", "ber3":
    "bertiga", "beresin": "membersihkan", "bete": "bosan"
```

Melakukan *cleaning teks* yaitu menghapus *mention*, *hashtag*, RT, link, angka, tanda baca dan karakter tidak penting.

cleaningText(text):

```
text = re.sub(r'@[A-Za-z0-9]+', "", text) # remove
mentions
text = re.sub(r'#[A-Za-z0-9]+', "", text) # remove
hashtag
text = re.sub(r'RT[\s]', "", text) # remove RT
text = re.sub(r"http\S+", "", text) # remove link
text = text.translate(str.maketrans("", string.punctuation)) #
remove all punctuations
text = text.strip()
return text
```

Membuat *case folding*: mengubah karakter menjadi huruf kecil, tokenisasi : memecah teks menjadi daftar

token, filterisasi : menghapus *stopword*, *Stemming* : mereduksi kata ke bentuk dasar, *remove consecutive duplicates* : menghapus kata duplikat yang bermunculan.

3.2 Pelabelan Data

Pelabelan data dengan menggunakan model BERT yang telah dilatih dengan label mapping: 0: negatif, 1: netral, 2: positif.

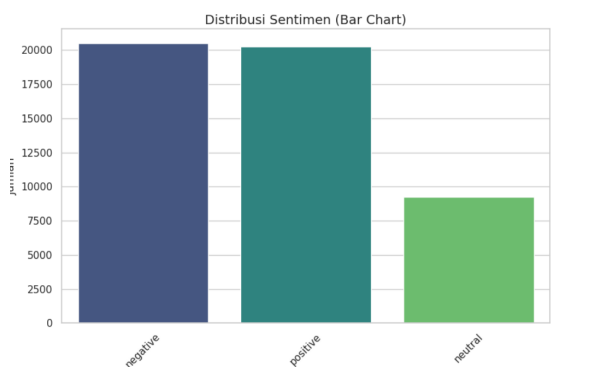
3.3 Distribusi Sentimen:

Hasil distribusi data ulasan / sentimen tersaji pada tabel 1.

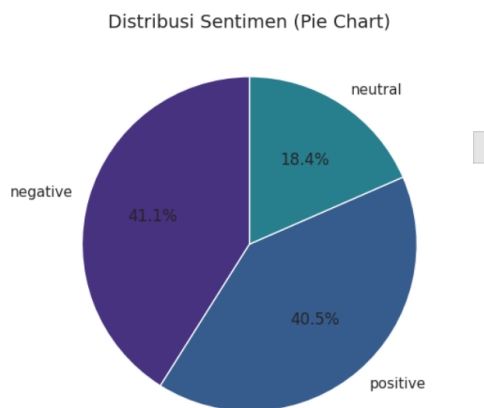
Tabel 1. Hasil Distribusi Sentimen

Sentimen	Jumlah
Negative	20527
Positive	20249
Neutral	9224

Kemudian data sentimen ditampilkan visualisasi distribusi data tersebut dalam *bar chart* dan *pie chart*



Gambar 2. Visualisasi Data Sentimen dalam Bar Chart



Gambar 3. Visualisasi Data Sentimen dalam Pie Chart

Dan terakhir dihitung presentase distribusi *sentiment* dan hasilnya pada tabel 2.

Tabel 2. Presentase Distribusi Sentimen

Presentase Distribusi Sentimen:	
Negative:	41.05%
Positive:	41.05%
Neutral:	18.45%

Ketidakeimbangan Data Sentimen

Ketidakeimbangan dalam data dapat memengaruhi kinerja model machine learning, terutama dalam tugas klasifikasi sentimen. Jika jumlah sampel untuk setiap kategori sentimen tidak seimbang, model cenderung lebih fokus pada kelas mayoritas (misalnya, sentimen netral) dan kurang sensitif terhadap kelas minoritas (misalnya, positif atau negatif). Akibatnya, meskipun akurasi keseluruhan tampak tinggi, model mungkin gagal mengenali kelas minoritas dengan baik. Grafik di atas menunjukkan ketidakeimbangan yang mencolok dalam distribusi sentimen, di mana jumlah komentar dengan sentimen netral secara signifikan lebih sedikit dibandingkan dengan sentimen positif dan negatif. Ketidakeimbangan ini dapat menyebabkan model kurang optimal dalam memprediksi kelas minoritas. Untuk mengatasi masalah ini, diperlukan penyesuaian agar data menjadi lebih seimbang. Salah satu pendekatan yang efektif untuk mengatasi ketidakeimbangan data adalah SMOTE (Synthetic Minority Over-sampling Technique). SMOTE digunakan untuk membuat keseimbangan jumlah data antara kelas mayoritas dan bukan mayoritas, bekerja dengan cara membuat data tambahan untuk kelas minoritas sehingga proporsi data antar kelas menjadi lebih merata.

Beberapa metode lain yang dapat digunakan meliputi: Pemberian Bobot Kelas: memberikan prioritas lebih tinggi pada kelas minoritas selama pelatihan model. Undersampling: mengurangi jumlah sampel dari kelas mayoritas untuk menciptakan keseimbangan. Oversampling: menambahkan lebih banyak sampel dari kelas minoritas, baik secara sintesis maupun dengan duplikasi. Dengan menerapkan salah satu atau kombinasi teknik di atas, diharapkan model dapat belajar dengan lebih baik dan memberikan hasil prediksi yang lebih adil serta akurat untuk semua kategori sentimen.

3.4 Pengujian Model

Model yang dibangun adalah LSTM + SMOTE + BERT, nilai epoch 5 dengan tahapan: Melakukan keseimbangan dataset dengan SMOTE, membuat model LSTM untuk mengurangi *overfitting* dengan tahapan: *bidirectional LSTM*, *Batch Normalization*, *Dropout*, *LSTM1*, *Batch Normalization_1* seperti tersaji pada tabel 3.

Tabel 3. Hasil Pengujian Model

Layer	Output Shape	Param #
bidirectional (Bidirectional)	(none, 1, 256)	918,528
batch_normalization (BatchNormalization)	(none, 1, 256)	1,024
dropout	(none, 1, 256)	0
Lstm_1(LSTM)	(none, 64)	82,176
batch_normalization_1 (batchnormalization)	(none, 64)	256
dropout_1 (dropout)	(none, 64)	0
Dense	(none, 3)	195

3.5 Evaluasi

Nilai akurasi untuk melatih model diperoleh nilai seperti pada tabel 4 dengan nilai batch 32 dan epoch 5.

Tabel 4. Nilai Akurasi Model Batch 32

Epoch	Waktu	Accuracy	Loss	Val_Accuracy	Val_Loss
1	17s	0,9157	1,4679	0,9898	0,9388
2	11s	0,9754	0,9360	0,9914	0,7095
3	11s	0,9808	0,7009	0,9937	0,5187
4	11s	0,9861	0,5109	0,9942	0,3742
5	11s	0,9868	0,3756	0,9931	0,2790

Sedangkan pada batch 64 dengan nilai epoch 5 diperoleh nilai seperti pada tabel 5.

Tabel 5. Nilai Akurasi Model 64

Epoch	Waktu	Accuracy	Loss	Val_Accuracy	Val_Loss
1	9s	0,9104	1,5616	0,9911	1,0164
2	6s	0,9774	1,0126	0,9902	0,8332
3	6s	0,9823	0,8256	0,9937	0,6757
4	6s	0,9858	0,6702	0,9965	0,5426
5	6s	0,9873	0,5407	0,9913	0,4461

Pada nilai batch 128 dan epoch 5, hasilnya seperti pada tabel 6.

Tabel 6. Nilai Akurasi Model 128

Epoch	Waktu	Accuracy	Loss	Val_Accuracy	Val_Loss
1	6s	0,9012	1,6439	0,9900	1,1362
2	3s	0,9804	1,1090	0,9912	0,9388
3	3s	0,9828	0,9301	0,9929	0,8063
4	3s	0,9849	0,8034	0,9895	0,7074
5	3s	0,9874	0,6950	0,9956	0,6042

Sedangkan pada saat melakukan data latih, hasil evaluasi training set pada batch 32 diperoleh nilai seperti pada tabel 7.

Tabel 7. Evaluasi Data Latih

Akurasi Training Set : 0.9954
Waktu : 3s 2ms/step

Sentimen	Precision	Recall	Fi-Score	Support
Negatif	0.99	1.00	0.99	12328
Neutral	0.99	1.00	1.00	14328
Positif	1.00	0.99	0.99	12328
Accuracy			1.00	42984

Untuk data latih, hasil evaluasi training set pada batch 64 didapat seperti pada tabel 8.

Tabel 8. Evaluasi Data Latih

Akurasi Training Set : 0.9942
Waktu : 3s 2ms/step

Sentimen	Precision	Recall	Fi-Score	Support
Negatif	1.00	0.99	0.99	14328
Neutral	0.99	1.00	0.99	14328
Positif	0.99	0.99	1.00	14328
Accuracy			0.99	42984

Untuk data latih, hasil evaluasi training set pada batch 128 diperoleh nilai seperti pada tabel 9.

Tabel 9. Evaluasi Data Latih batch 128

Akurasi Training Set : 0.9967
Waktu : 3s 2ms/step

Sentimen	Precision	Recall	Fi-Score	Support
Negatif	1.00	0.99	1.00	14328
Neutral	0.99	1.00	1.00	14328
Positif	1.00	1.00	1.00	14328
Accuracy			1.00	42984

Dan evaluasi untuk data tes pada batch 32 diperoleh nilai seperti pada tabel 10.

Tabel 10. Evaluasi Hasil Data Tes

Akurasi Test Set : 0.9927
Waktu : 1s 2ms/step

Sentimen	Precision	Recall	Fi-Score	Support
Negatif	0.99	1.00	0.99	6199
Neutral	0.98	1.00	0.99	2724
Positif	1.00	0.99	0.99	6077
Accuracy			0.99	15000

Dan evaluasi untuk data tes pada batch 64 diperoleh nilai seperti pada tabel 11.

Tabel 11. Evaluasi Hasil Data Tes

Akurasi Tes Set : 0.9927
Waktu : 1s 2ms/step

Sentimen	Precision	Recall	Fi-Score	Support
Negatif	0.99	1.00	0.99	6199
Neutral	0.98	1.00	0.99	2724
Positif	1.00	0.99	0.99	6077
Accuracy			0.99	15000

Dan evaluasi untuk data tes pada batch 128 didapat seperti pada tabel 12.

Tabel 12. Evaluasi Hasil Data Tes

Akurasi Tes Set : 0.9956
Waktu : 1s 2ms/step

Sentimen	Precision	Recall	Fi-Score	Support
Negatif	1.00	0.99	1.00	6199
Neutral	0.98	1.00	0.99	2724
Positif	1.00	1.00	1.00	6077
Accuracy			1.00	15000

4. Kesimpulan

Berdasarkan hasil pengujian LSTM dengan menggunakan batch 32, 64 dan 128 dan nilai *epoch* = 5 diperoleh hasil terbaik pada batch 128 mampu memberikan performa terbaik pada nilai akurasi sebesar 99.56%. Nilai precision pada model LSTM menunjukkan hasil terbaik pada sentimen negatif sebesar 1.00, sentimen positif sebesar 1.00 dan nilai *recall* sebesar 0.99. Penerapan LSTM dengan SMOTE dan BERT memberikan performa terbaik untuk analisis sentimen ulasan penggunaan aplikasi Shopee dari *Google play Store*.

Daftar Rujukan

- [1] S. Simatupang *et al*, "Facebook Marketplace Serta Pengaruhnya Terhadap Minat Beli", Jurnal Ekbis Volume 22 No 1, pp 28-41, 2021.

-
- [2] R. Yustiani and R. Yunanto, "Peran Marketplace Sebagai Alternatif Bisnis Di Era Teknologi Informasi", *Jurnal Ilmiah Komputer dan Informatika (KOMPUTA)* Vol 6(2), pp. 43–48, 2017.
- [3] E. Y. Nasution *et al*, "Perkembangan Transaksi Bisnis E-Commerce terhadap Pertumbuhan Ekonomi di Indonesia," *Jesya*, vol. 3, no. 2, pp. 506–519, 2020.
- [4] I. D. Safitri, "Peran marketplace dalam meningkatkan ekonomi umat: studi pada Ali Dien Marketplace Surabaya" (Doctoral dissertation, UIN Sunan Ampel Surabaya), 2020.
- [5] J. Sintadi and Y. Yoestini, "Analisis Pengaruh Brand Image, Harga, Dan Review Produk Terhadap Keputusan Pembelian Smartphone Samsung Pada Mahasiswa Undip Semarang," *Diponegoro Journal of Management*, vol. 8, no. 2, pp. 33–43, 2019.
- [6] R.M Arrasyid *et al*, "Analisis Sentimen Review Pembelian di Marketplace Shopee Menggunakan Pendekatan Natural Language Processing", *Jurnal Tekno KOMPAK* Volume 18 No. 2, pp.319-330, 2024.
- [7] G. S. Lasatira *et al*, "Analisis Sentimen Terhadap Pengaruh Minat Belanja Berdasarkan Komentar di Marketplace Menggunakan Metode Recurrent Neural Network (RNN)". *Jurnal Sistem Informasi Bisnis* 02(2023) pp. 112 – 119, 2023.
- [8] K. Ivanedra and M. Mustikasari, "Implementasi Metode Recurrent Neural Network Pada Text Summarization Dengan Teknik Abstraktif", *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 6(4), pp. 377 – 382, 2019.
- [9] M. Musfiroh *et al*, "Analisis Sentimen Terhadap Ulasan Aplikasi Shopee di Google Play Store Menggunakan Metode TF-IDF dan Long Short-Term Memory (LSTM)". *Journal of Electrical Engineering and Computer (JEECOM)* Vol. 6, No. 2 pp. 371 – 381, 2024.
- [10] N. Selle, *et al*, "Perbandingan Prediksi Penggunaan Listrik Dengan Menggunakan Metode Long Short Term Memory (LSTM) Dan Recurrent Neural Network (RNN)". *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)* Volume 9 No. 1, pp. 155 – 162, 2022.
- [11] A. Anggito and J. Setiawan, *Metodologi Penelitian Kualitatif*. Jawa Barat: CV Jejak, 2018.
- [12] W. Wartumi and Y.A. Wijaya, "Analisis Data Sentimen Ulasan Pengguna Aplikasi Shopee di Google Play Store dengan Klasifikasi Algoritma Naïve Bayes". *Jurnal Informatika dan Rekayasa Perangkat Lunak* Vol. 6, No. 1, pp. 164-170, 2024.
- [13] D. Septiani and I. Isabela, "Analisis Term Frequency Inverse Document Frequency (TF-IDF) Dalam Temu Kembali Informasi Pada Dokumen Teks". *Jurnal Sistem dan Teknologi Informasi Indonesia* Vol 1 Nomor 2, pp 81-88, 2022.
- [14] Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press. <https://doi.org/10.1017/9781108638510>
- [15] Zhang, Y., & Wallace, B. C. (2017). A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. In *Proceedings of the 8th International Joint Conference on Natural Language Processing* (pp. 253–263).