

ANALISIS KEPUASAN PELANGGAN PADA PELAYANAN E-COMMERCE TOKOPEDIA DAN SHOPEE PADA TWITTER MENGGUNAKAN RECURRENT NEURAL NETWORK

Syahrul Alfani¹⁾, Rudi Hariyanto²⁾, Nanda Martyan Anggadimas³⁾

Universitas Merdeka Pasuruan
syahrulalfani@student.unmerpas.ac.id

Abstrak

Salah satu media sosial yang mempengaruhi transaksi e-commerce tersebut adalah Twitter, karena salah satu media sosial yang sangat populer di Indonesia. Menurut Kementerian Komunikasi dan Informatika, Data pengguna Twitter ini dapat digunakan untuk menganalisis perkembangan sistem yang lebih baik dalam e-commerce, dan analisis yang dilakukan akan berguna untuk mengetahui pendapat pengguna Twitter tentang efisiensi pelanggan e-commerce. Proyek yang dipilih dalam penelitian ini adalah platform e-commerce dari perusahaan teknologi Indonesia Tokopedia dan perusahaan teknologi asing Shopee. Recurrent Neural Network(RNN) adalah bagian dari jaringan saraf untuk memproses data yang terhubung (data berurutan). Algoritma RNN digunakan dalam analisis ini, yang menunjukkan kinerja yang baik dalam memproses data seperti teks. Identifikasi permasalahan yang ada yaitu persaingan antara e-commerce Tokopedia dan Shopee yang terdapat pada ulasan/tweets pengguna di sosial media Twitter. Membandingkan kepuasan pelayanan yang dirasakan pelanggan terhadap kedua e-commerce menggunakan Recurrent Neural Network (RNN) untuk mendapatkan nilai akurasi yang terbaik. Pencarian literatur berdasarkan topik penelitian yang telah ditentukan pada tahap awal penelitian ini. ulasan/tweets pengguna di sosial media Twitter dijadikan sebagai sumber data yang digunakan dalam penelitian untuk menyeleksi variabel. Confusion matrix digunakan untuk mengevaluasi kinerja sebuah model klasifikasi dengan membandingkan label asli dan prediksi label dari set data yang diberikan. Pada kasus ini, kita memiliki *confusion matrix* dengan dua kelas yang direpresentasikan sebagai 0 (*negative*) dan 1 (*positive*). Berdasarkan pembahasan dan hasil analisis, bahwa ulasan tweets yang berjumlah 200 data Tokopedia dan 200 data Shopee. Dalam proses pengujian klasifikasi dari hasil perhitungan *confusion matrix* yang menjadi acuan untuk prediksi dan aktual, lalu menggunakan model algoritma yaitu Simple RNN. Pada ulasan pelanggan lebih banyak menyukai pelayanan Tokopedia dari user Twitter dapat dilihat dengan data uji yang positif sebesar 140 data sedangkan data uji yang positif pada pelayanan Shopee sebesar 120 data. Hasil akurasi yang memperoleh performa lebih baik yaitu Tokopedia menggunakan algoritme Simple RNN dengan nilai akurasi mencapai 53%, presisi mencapai 57%, recall mencapai 50%, dan f1-score mencapai 53,27%. Jika dibandingkan nilai klasifikasi terbaik dari data Shopee dari algoritme Simple RNN menunjukkan nilai klasifikasi dengan hasil akurasi sebesar 47%, presisi mencapai 60%, recall sebesar 33%, dan f1-score mencapai 40%.

Kata Kunci : E-Commerce; Kepuasan Pelanggan; Jaringan Saraf Tiruan; Recurrent Neural Network, Twitter.

Abstract

One of the social media that influences e-commerce transactions is Twitter, because it is one of the most popular social media in Indonesia. According to the Ministry of Communication and Information, Twitter user data can be used to analyze the development of better systems in

Analisis Kepuasan Pelanggan Pada Pelayanan

E-Commerce Tokopedia Dan Shopee Pada Twitter Menggunakan Recurrent Neural Network

e-commerce, and the analysis carried out will be useful to find out Twitter users' opinions about e-commerce customer efficiency. The project chosen in this research is a platform e-commerce from Indonesian technology company Tokopedia and foreign technology company Shopee. Recurrent Neural Network (RNN) is a part of a neural network for processing connected data (sequential data). The RNN algorithm is used in this analysis, which shows good performance in processing data such as text. Identify existing problems, namely competition between e-commerce Tokopedia and Shopee which can be found in user reviews/tweets on Twitter social media. Comparing the service satisfaction felt by customers for both e-commerce sites using Recurrent Neural Network (RNN) to get the best accuracy value. Literature search based on research topics that have been determined at the initial stage of this research. User reviews/tweets on Twitter social media are used as a data source used in research to select variables. Confusion matrix is used to evaluate the performance of a classification model by comparing original labels and predicted labels from the given data set. In this case, we have a confusion matrix with two classes represented as 0 (negative) and 1 (positive). Based on the discussion and analysis results, the tweet reviews amount to 200 Tokopedia data and 200 Shopee data. In the classification testing process, the results of the confusion matrix calculation are used as a reference for predictions and actuals, then use an algorithm model, namely Simple RNN. In customer reviews, Twitter users prefer Tokopedia services, which can be seen from the positive test data of 140 data, while the positive test data for Shopee services is 120 data. The accuracy results that obtained better performance were Tokopedia using the Simple RNN algorithm with an accuracy value of 53%, precision of 57%, recall of 50%, and f1-score of 57%. When compared to the best classification value from Shopee data from the Simple RNN algorithm, it shows a classification value with accuracy results of 47%, precision reaching 60%, recall of 33%, and f1-score reaching 40%.

Kata Kunci : E-Commerce, Customer Satisfaction, Artificial Neural Networks, Recurrent Neural Network, Twitter.

1. PENDAHULUAN

Dalam perkembangannya, internet bukan hanya digunakan sebagai media informasi dan komunikasi saja, adapun dampak pengguna internet, yaitu penggunaan media sosial yang terus meningkat (Kautsar,2018). Perkembangan dunia internet juga mempengaruhi proses transaksi jual dan beli yang awalnya bersifat konvensional berubah menjadi *modern* dengan bantuan internet, yang disebut juga sebagai *e-commerce*. Salah satu media sosial yang mempengaruhi transaksi *e-commerce* tersebut adalah Twitter, karena salah satu media sosial yang sangat populer di Indonesia. Menurut Kementerian Komunikasi dan Informatika (Kemenkominfo), Indonesia menempati peringkat 5 terbanyak pengguna Twitter terbesar di dunia. Posisi Indonesia hanya kalah dari USA, Brazil, Jepang, dan Inggris (Kominfo, 2019).

Data pengguna Twitter ini dapat digunakan untuk menganalisis perkembangan sistem yang lebih baik dalam *e-commerce*, dan analisis yang dilakukan akan berguna untuk mengetahui pendapat pengguna Twitter tentang efisiensi pelanggan *e-commerce*. Saat ini banyak sekali perkembangan *e-commerce* yang terjadi di Indonesia, seperti Shopee, Lazada, Tokopedia dan Bukalapak. Proyek yang dipilih dalam penelitian ini adalah platform *e-commerce* dari perusahaan teknologi Indonesia Tokopedia dan perusahaan teknologi asing Shopee. Karena jejaring sosial seperti Twitter berperan penting dalam mempengaruhi aspek sosial, komunikatif, psikologis, pemasaran dan politik, maka data yang diperoleh untuk penelitian ini diperoleh dari Twitter.

2. METODE

2.1 Identifikasi Masalah

Identifikasi permasalahan yang ada yaitu persaingan antara *e-commerce* Tokopedia dan Shopee yang terdapat pada ulasan/tweets pengguna di sosial media *Twitter*. Membandingkan kepuasan pelayanan yang dirasakan pelanggan terhadap kedua *e-commerce* menggunakan *Recurrent Neural Network* (RNN) untuk mendapatkan nilai akurasi yang terbaik..

2.2 Analisis Kebutuhan Sistem

Penelitian ini dilakukan dengan menggunakan alat untuk mendukung dan menunjang pelaksanaan penelitian, yaitu sebagai berikut:

2.2.1 Perangkat keras (*Hardware*)

Berikut adalah Spesifikasi Perangkat keras yang digunakan penulis pada penelitian ini.

Tabel 1 Hardware

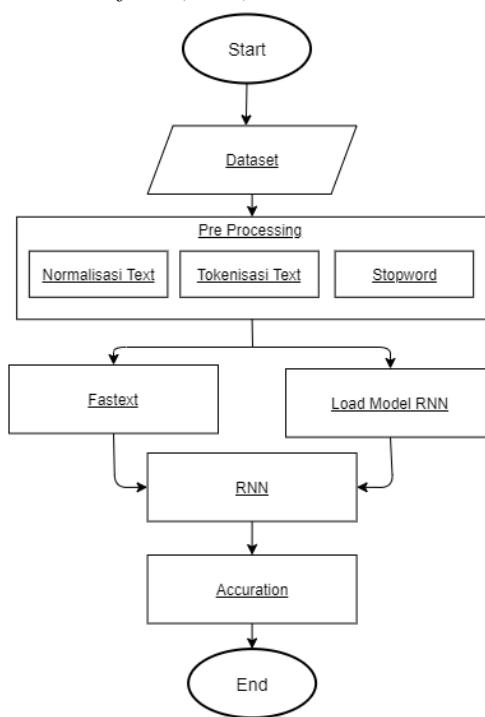
No.	Nama	Spesifikasi
1.	Laptop	Intel Celeron 2,16 Ghz, Memory 4 GB, SSD 120 GB
2.	OS	Windows 10

3.2.1 Perangkat Lunak (*Software*)

Pada penelitian ini penulis menggunakan *Microsoft Excel*, *Google Colaboration* dan *Google Chrome*. Untuk *Microsoft Excel* digunakan untuk *dataset*, *Google Colaboration* digunakan untuk *fasttext*, *preprocessing*, dan metode, dan Untuk *Google Chrome* digunakan untuk menampilkan *Graphical User interface*(GUI).

2.3 Desain Sistem

2.3.1 Flowchart Sistem



Gambar 1 Flowchart Sistem Training dan Testing

Algoritma dari sistem yang akan digunakan untuk menganalisis opini keluhan dan kecewa dari *user Twitter* atas pelayanan kedua *e-commerce* dan menunjukkan opini kepuasan dan kesenangan atas pelayanan kedua *e-commerce* dengan metode *Recurrent Neural Network* (RNN) ini akan dibangun seperti pada gambar 1 *Flowchart Sistem Training dan Testing*.

2.3.2 Dataset

Dataset yang berupa data ulasan kepuasan pelayanan *e-commerce* Tokopedia dan Shopee pada media sosial Twitter dengan mengumpulkan *tweet* mengenai bagaimana pelayanan kedua *e-commerce* tersebut.

2.4.3 Preprocessing

Preprocessing adalah metode awal untuk mengubah data mentah menjadi informasi dan format yang lebih efektif dan berguna, dikenal sebagai data *preprocessing*. Format *raw* data dari berbagai sumber sering mengalami *error*, *missing value*, dan *inkonsistensi*. Akibatnya, format perlu diubah agar hasil data mining tepat(Sani & Sarwani,2022). Tahap *preprocessing* dengan menggunakan metode *Recurrent Neural Network*(RNN) yang bertujuan untuk menerapkan analisis data pada *Twitter* untuk isu *marketing*. Dan, untuk Penelitian ini *preprocessing* data terdapat tiga proses yaitu :

a. Normalisasi Teks

Normalisasi teks berfungsi untuk melakukan penormalan kata atau kalimat dalam suatu data, normalisasi akan dilakukan pada seluruh data teks penelitian yang digunakan. Proses normalisasi teks akan menghilangkan sekumpulan simbol-simbol seperti seperti tanda titik, tanda koma, tanda tanya, tanda seru dan simbol lainnya atau URL (Muludi et al., 2021).Contoh :

Tabel 2 Normalisasi Teks

Data Kalimat	Proses Normalisasi Teks
@maspiyuu Cuma satu orang yg bisa ngalahin Jokowi yaitu anies baswedan. Tapi karna Prabowo ngotot, yah kemungkinan??\$?https://t.co/vhk4aw0jhk	Cuma satu orang yang bisa mengalahkan Jokowi yaitu anies baswedan. Tetapi karena Prabowo keras kepala kemungkinan

b. Tokenisasi Teks

Tahap ini merupakan tahap pemotongan *string* atau kalimat *input* berdasarkan tiap kata yang menyusun kalimat dalam teks. Tokenisasi secara garis besar memecah sekumpulan karakter dalam suatu teks ke dalam satuan kata, bekerja dengan membedakan karakter-karakter tertentu yang dapat diperlakukan sebagai pemisah kata atau bukan sebuah kata (Muludi etal., 2021). Contoh :

Tabel 3 Tokenisasi Teks

Data Kalimat	Proses Tokenisasi Teks
Cuma satu orang yg bisa ngalahin Jokowi yaitu anies baswedan. Tapi karna Prabowo ngotot,yah Kemungkinan	[‘Cuma’, ‘Satu’, ‘Orang’, ‘Yg’, ‘bisa’, ‘Ngalahin’, ‘Jokowi’, ‘Yaitu’, ‘Anies’, ‘Baswedan’, ‘Tapi’, ‘Karna’, ‘Prabowo’, ‘Ngotot’, ‘Ya’, ‘Kemungkinan’, ‘.’]

c. Stopword

Stopword adalah sekumpulan kata-kata yang sering muncul tetapi jika dihapus tidak mengubah makna dari kalimat didalam dataset yang diperoleh dari *tweet* tersebut. Pembuangan *stopword* dilakukan untuk mengetahui suatu kata yang tidak memiliki arti. Kata yang telah diproses melalui tahap tokenisasi diperiksa dalam *stopword*, apabila ada sebuah kata masuk di dalam *stopword* maka kata itu tidak akan diproses lebihlanjut (Muludi et al., 2021). Contoh :

Tabel 4 Stopword

Data Kalimat	Proses Stopword
Cuma satu orang yg bisa ngalahin Jokowi yaitu anies baswedan. Tapi karna Prabowo ngotot,yah kemungkinan	[‘’, ‘Cuma’, ‘Satu’, ‘Orang’, ‘Yang’ , ‘Bisa’, ‘Ngalahin’, ‘Jokowi’, ‘Yaitu’, ‘Anies’, ‘Baswedan’, ‘.’, ‘Tapi’, ‘Karna’, ‘Prabowo’, ‘Ngotot’, ‘ya’, ‘kemungkinan’, ‘.’]

2.4.4 Fasttext

Proses *fasttext* untuk merepresentasi kata menjadi vector dimulai dari *input* (x) token kata yang didapat dari hasil proses *preprocessing* data, lalu dikalikan dengan bobot untuk mendapatkan nilai *hidden* (h). jadi nilai *hidden* (h) tersebut yang akan dijadikan vektor. Dengan menggunakan lima *window parameter*, kami dapat memperoleh dataset dengan ukuran vektor 300 untuk penelitian ini. Proses perhitungan diamati dari lima kata yang bertetangga, baik sebelum maupun sesudahnya. Contoh kata demi kata dari nilai vektor ditunjukkan pada table dibawah ini.

Tabel 5 Hasil Vektor

get_word_vector	Word	Vector
	Cuma	0,24
	Satu	0,23
	Bisa	0,22
Capres	Yg	0,21

3. HASIL DAN PEMBAHASAN**3.1 IMPLEMENTASI SISTEM**

Implementasi sistem ialah hasil penelitian pada sistem yang telah diciptakan. Berikut merupakan beberapa *interface* yang terdapat dalam sistem.

```
[9] import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from tensorflow.keras.preprocessing.text import Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import SimpleRNN, Embedding, Dense, Dropout, Bidirectional
from tabulate import tabulate
import fasttext
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import confusion_matrix
import re
import nltk
from nltk.corpus import stopwords
from tensorflow.keras.preprocessing.text import Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
from flask import Flask, render_template, request
from flask_ngrok import run_with_ngrok
from flask import Flask, render_template
from pyngrok import ngrok
```

Gambar 2 Tampilan Library

Yang dilakukan pertama adalah mengimport *library* yang dibutuhkan ada beberapa yang harus *diimport* adalah *numpy, pandas, sklearn.model_selection, sklearn.preprocessing, tensorflow.keras.preprocessing.text, tensorflow.keras.preprocessing.sequence, tensorflow.keras.models, tensorflow.tensorflow.keras.layers, tabulate, sklearn.metrics, nltk.corpus, ngrok, dan fasttext*.

Selanjutnya melakukan *load model fasttext* untuk model Wikipedia Indonesia.

```
[10] # Load FastText model
fasttext_model = fasttext.load_model('cc.id.300.bin')
# For older versions of fasttext, you might need to use the following:
# fasttext_model = fasttext.load_model('cc.id.300.bin', label_prefix='_label_')
```

Gambar 3 Load Model Fasttext

```
import pandas as pd
from google.colab import files

# Upload the Excel file
uploaded = files.upload()

# Get the uploaded file name
file_name = list(uploaded.keys())[0]

# Import Dataset
df = pd.read_excel(file_name)
df.head(10)
```

[Plain File] dataset_shopee.xls
dataset_shopee.xls(application/vnd.ms-excel) - 39424 bytes, last modified: 3/2/2024 - 100% done
Saving dataset_shopee.xls to dataset_shopee (1).xls

Nomor	username	komentar	Unnamed: 3	rating	kategori
0	xasyera	tassynp convomfs halii kak izin reply ya aku p...	NaN	1 star	Negative
1	anytimeunneed_	dari tanggal dianggurin baru tgl dibales cob...	NaN	1 star	Negative
2	RaodaKayana	Hi kak maaf ketidaknyamanannya terkait kendala...	NaN	1 star	Negative
3	sweetiepledenu	Shopee express pelayanan mu jelek banget ya su...	NaN	1 star	Negative
4	alpakinr	AnterajaCare alpakinr hi kak maaf buat resah ...	NaN	1 star	Negative
5	eskimcookyat	Dear X please do ur magic Sangat kecewa dgn pe...	NaN	5 stars	Positive
6					
7					
8					
9					
10					

Gambar 4 Import Dataset

Kemudian dilakukan *import dataset* kedalam program. *Dataset* pada penelitian ini dibuat dalam bentuk *excel* dan dipanggil kedalam sistem yaitu “*df = pd.read_excel(file_name)*. (*file_name*) adalah nama dari file excel yang disimpan sesuai dengan *name type* dari file tersebut Output yang ditampilkan pada gambar tersebut adalah 10 baris, diawali dari angka 0 sampai 9.

```
import pandas as pd

# Assuming 'df' is the DataFrame containing the 'kategori' column

# Add the "_label_" prefix to the 'kategori' column
df['kategori'] = '_label_' + df['kategori'].astype(str)

# Save the DataFrame to a CSV file
df.to_csv('dataset_fasttext.csv', index=False)
df.head(10)
```

Nomor	username	komentar	Unnamed: 3	rating	kategori
0	xasyera	tassynp convomfs halii kak izin reply ya aku p...	NaN	1 star	_label_Negative
1	anytimeunneed_	dari tanggal dianggurin baru tgl dibales cob...	NaN	1 star	_label_Negative
2	RaodaKayana	Hi kak maaf ketidaknyamanannya terkait kendala...	NaN	1 star	_label_Negative
3	sweetiepledenu	Shopee express pelayanan mu jelek banget ya su...	NaN	1 star	_label_Negative
4	alpakinr	AnterajaCare alpakinr hi kak maaf buat resah ...	NaN	1 star	_label_Negative
5	eskimcookyat	Dear X please do ur magic Sangat kecewa dgn pe...	NaN	5 stars	_label_Positive
6	l_a_paw_logize	Teryata balikin barang buat dikirim ulang lebi...	NaN	1 star	_label_Negative
7	keigra	Mohon pastikan kakak tetap melakukan pembayar...	NaN	5 stars	_label_Positive
8	zxyaan	siopatekspres zxyaan Hi Kak terkait kendala ...	NaN	1 star	_label_Negative
9	caisyr	Hi kak mohon maaf atas kendala yg dialami Sho...	NaN	1 star	_label_Negative
10					

Gambar 5 Pretrained Dataset dengan fasttext

Dilakukan pada *DataFrame* yang disimpan dalam variable ‘*df*’ yang memiliki kolom ‘*kategori*’. Untuk setiap nilai dalam kolom’*kategori*’, kode tersebut menambahkan awalan “*label*” dengan menggabungkan string “*label* dengan setiap nilai menggunakan operator “+”. Misalnya, jika ada nilai “A” dalam ‘*kategori*’, Setelah diubah dengan kode tersebut, nilainya akan menjadi “*_label_A*”.

```

# Membaca dataset
dataset = pd.read_csv('dataset_fasttext.csv')

# Pra-pemrosesan
X = dataset['komentar'].values
y = dataset['kategori'].values

# Negerjemahkan label
label_encoder = LabelEncoder()
y = label_encoder.fit_transform(y)

# Membagi dataset menjadi data latih dan data uji
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Menyimpan data uji ke dalam file CSV
test_data = pd.DataFrame({'komentar': X_test, 'kategori': y_test})
test_data.to_csv('test_data.csv', index=False)

# Menyimpan data latih ke dalam file CSV
train_data = pd.DataFrame({'komentar': X_train, 'kategori': y_train})
train_data.to_csv('train_data.csv', index=False)

# Mencetak bentuk (shape) dari data uji dan data latih
print("Bentuk data uji:", test_data.shape)
print("Bentuk data latih:", train_data.shape)

Bentuk data uji: (48, 2)
Bentuk data latih: (160, 2)

```

Gambar 6 Proses pembagian data

Dilakukan pembagian data menjadi data test dan data training. Pertama, panggil dataset terbaru yang tadi telah disave di proses Pretrained dalam bentuk csv. Lalu dilakukan pra-pemrosesan ('X = komplain' dan 'Y=kategori').

```

# Tokenization
tokenizer = Tokenizer(num_words=5000)

# Membersihkan karakter yang tidak penting dan menghapus stopwords sebelum tokenisasi
stop_words = set(stopwords.words('indonesian'))
X_train = [' '.join([word for word in re.sub(r'[^\w\s]', '', text.lower()).split() if word not in stop_words]) for text in X_train]
X_test = [' '.join([word for word in re.sub(r'[^\w\s]', '', text.lower()).split() if word not in stop_words]) for text in X_test]

tokenizer.fit_on_texts(X_train)

X_train = tokenizer.texts_to_sequences(X_train)
X_test = tokenizer.texts_to_sequences(X_test)

```

Gambar 7 Tokenizing dan StopWord

Selanjutnya, pada tahapan pra-pemrosesan data teks, dilakukan tokenisasi dengan menggunakan objek ‘Tokenizer’ dari Keras dengan batasan jumlah kata maksimal sebanyak 5000.

```

max_length = 100 # Panjang maksimal sekuen kata yang akan dipertimbangkan

X_train = pad_sequences(X_train, maxlen=max_length, padding='post')
X_test = pad_sequences(X_test, maxlen=max_length, padding='post')

```

Gambar 8 Proses Padding

Pada Tahap ini ‘pad_squences’ dari modul ‘tensorflow.keras.preprocessing.squence’ digunakan untuk melakukan paddin pada urutan bilangan bulat yang telah dihasilkan sebelumnya. Panjang maksimal sekuen kata yang akan dipertimbangkan adalah 100, dan padding dilakukan pada akhir sekuen kata menggunakan nilai ‘padding=”post”’.

```

# membuat word embedding dengan model fasttext
embedding_dim = 300
word_index = tokenizer.word_index
embedding_matrix = np.zeros((len(word_index) + 1, embedding_dim))
for word, i in word_index.items():
    if word in fasttext_model:
        embedding_matrix[i] = fasttext_model[word]

```

Gambar 9 Proses word embedding dengan model fasttext

Pada tahap ini digunakan untuk membuat *word embedding* dengan model *FastText*. Dimana dimensi *embedding* adalah 300, dan *vector* representasi kata-kata akan disimpan dalam matriks’*embedding_matrix*’ dengan menggunakan model *FastText* yang diberikan (‘*fasttext_model*’).

```

# Build the model
model = Sequential()
model.add(Embedding(len(tokenizer.word_index) + 1, 300, input_length=max_length))
model.add(SimpleRNN(128))
model.add(Dense(1, activation='sigmoid'))

model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])
model.summary()

```

Gambar 10 Permodelan RNN Tabel

Pada tahap ini, Model yang didefinisikan adalah *Sequentaial model*, yang berarti layer-layer berurutan.

```

# Training the model
epochs = 25
batch_size = 64
history = model.fit(X_train, y_train, epochs=epochs, batch_size=batch_size, validation_data=(X_test, y_test))

# Save the model
model.save('sentiment_model.h5')

Epoch 1/25
3/3 [=====] - 3s 390ms/step - loss: 1.0029 - accuracy: 0.6125 - val_loss: 0.8252 - val_accuracy: 0.7000
Epoch 2/25
3/3 [=====] - 0s 100ms/step - loss: 1.0194 - accuracy: 0.5688 - val_loss: 0.7737 - val_accuracy: 0.7000
Epoch 3/25
3/3 [=====] - 0s 77ms/step - loss: 0.7625 - accuracy: 0.7063 - val_loss: 0.7887 - val_accuracy: 0.7000
Epoch 4/25
3/3 [=====] - 0s 79ms/step - loss: 0.7497 - accuracy: 0.7125 - val_loss: 0.7868 - val_accuracy: 0.7000
Epoch 5/25
3/3 [=====] - 0s 82ms/step - loss: 0.7225 - accuracy: 0.7003 - val_loss: 0.7085 - val_accuracy: 0.7000
Epoch 6/25
3/3 [=====] - 0s 87ms/step - loss: 0.6981 - accuracy: 0.7063 - val_loss: 0.6733 - val_accuracy: 0.7000
Epoch 7/25
3/3 [=====] - 0s 78ms/step - loss: 0.6828 - accuracy: 0.7188 - val_loss: 0.7881 - val_accuracy: 0.6750
Epoch 8/25
3/3 [=====] - 0s 77ms/step - loss: 0.6473 - accuracy: 0.7250 - val_loss: 0.6208 - val_accuracy: 0.7500
Epoch 9/25
3/3 [=====] - 0s 84ms/step - loss: 0.5965 - accuracy: 0.7688 - val_loss: 0.7009 - val_accuracy: 0.7250
Epoch 10/25
3/3 [=====] - 0s 82ms/step - loss: 0.5263 - accuracy: 0.7563 - val_loss: 0.6648 - val_accuracy: 0.7250
Epoch 11/25
3/3 [=====] - 0s 79ms/step - loss: 0.5282 - accuracy: 0.7758 - val_loss: 0.6444 - val_accuracy: 0.7500

```

Gambar 11 Proses Training

Pada tahap ini, proses training dengan jumlah epochs 25 dan dengan *batchsize* 64.

```

import pandas as pd
import numpy as np

# Load the test data from the CSV file
test_data = pd.read_csv('test_data.csv')

# Get the 'komplain' column from the test data
X_test = test_data['komenter'].values

# Tokenization and padding (similar to the preprocessing steps before)
X_test = tokenizer.texts_to_sequences(X_test)
X_test = pad_sequences(X_test, maxlen=max_length, padding='post')

```

Gambar 12 Proses klasifikasi data testing

Pada tahap ini, Setelah melakukan import pustaka Pandas sebagai “pd” dan pustaka NumPy sebagai “np” data nii dimuat dari file CSV menggunakan ‘`nd read_csv()`’

enak amat shopee malay, ga
Shopee selalu memberikan informasi y
Diarahin mulu, emang gue tukang parkir!!!! Heran Shopee udah mah kejadian paket stuck tuh bukan sekali dua kali. Masih aja j
@thaanderstorm ANJ BGT ONGIRNYA 🤦 di shopee
Kurir shopee express pada kemana sih paket gue gak diantar2 sampe sekarang, dahlah gak usah diantar merajuk gue, sudah bel
Pelayanan pengiriman di S
Pelayanan Ø bgt deh shopee bener bener dikomplain blm selesai diakhiri chatnya wkwkwk SUMPAH GA T
Dalam kurun waktu kurang dari sebulan udah dikecawain sama pelayanan SPX 2 KALI 🤬 Please yg mau pake SPX mikir-mikir dulu deh,
Shopee memberikan pelayanan yang rama
Pelayanan pelanggan Shopee
Pelayanan gramedia official shopee jelek bgt sumpah. Diawal udah gue kasih list buku yg gue mau ada apa engga buat ngehind
@txtdarionlshop Ane kalo check out sopi, tak cek lagi Buat ganti ke courier lain selain @ShopeeXpress Lebih suka ke @jntexpressi
Ini bakal nimbulin ekspektasi kalau chat ke cs shopee bakal dijawab dengan hal yang serupa, padahal belum tentu semu
harus kuakui, pelayanan shopee xpress lebih baik dibanding jr
@fbry_hvk @JNECare @ShopeeID Benar, ini untuk tiktok udah mulai kerjasama dengan tokped. Saya mulai meninggalkan shopee pelan2. Kl tidak d

Gambar 13 Hasil klasifikasi data testing

Gambar 13 menunjukkan bahwa hasil pengklasifikasi sistem lebih baik dari pada klasifikasi manual.

```

# Print accuracy for training and testing
print(f'Akurasi Training: {history.history["accuracy"][-1]:.2f}')
print(f'Akurasi Testing: {history.history["val_accuracy"][-1]:.2f}')

Akurasi Training: 1.00
Akurasi Testing: 0.47

```

Gambar 14 Hasil akurasi Training dan Testing Shopee

Gambar 14 menunjukkan hasil akurasi training dan testing. Dimana akurasi training lebih tinggi dibandingkan testing.

- a. Akurasi (*accuracy*) *presentase* dari jumlah total prediksi yang benar.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\%$$

$$\text{Accuracy} = \frac{3 + 4}{3 + 4 + 6 + 2} \times 100\% = \frac{7}{15} \times 100\% = 47\% / 0,47$$

- b. Presisi (*precision*) adalah porsi kasus positif yang teridentifikasi dengan benar. $\underline{= 33\% / 0,3}$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\%$$

$$\text{Precision} = \frac{3}{3 + 6} \times 100\% = \frac{3}{9} \times 100\%$$

- c. *Recall* merupakan *presentase* kasus positif yang teridentifikasi dengan benar.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\%$$

$$= 60\% / 0,6$$

$$\text{Recall} = \frac{3}{3 + 2} \times 100\% = \frac{3}{5} \times 100\%$$

- d. *F1-Score* merupakan perbandingan rata-rata presisi dan *recall* yang dibobot.

$$f1 score = 2 \times \frac{(\text{recall} \times \text{precision})}{(\text{recall} + \text{precision})} \times 100\%$$

$$f1 score = 2 \times \frac{(0,3 \times 0,6)}{(0,3 + 0,6)} \times 100\%$$

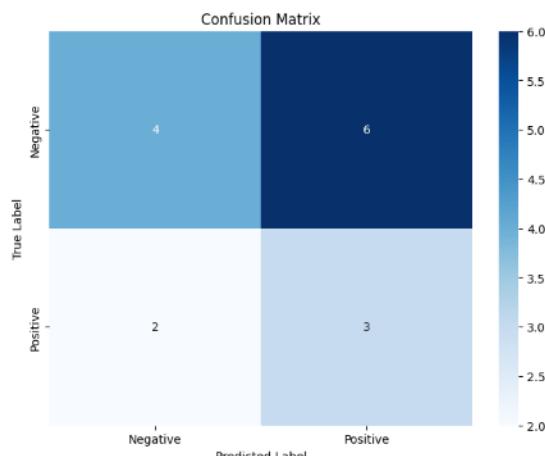
$$f1 score = 2 \times \frac{0,18}{0,9} \times 100\% = 40\%$$

Keterangan : True Positive (TP) = 3

True Negative (TN) = 4

False Positive(FP) = 6

False Negative(FN) = 2



Gambar 15 Hasil RNN Shopee dalam Confusion Matrix

Confusion matrix digunakan untuk mengevaluasi kinerja sebuah model klasifikasi dengan membandingkan label asli dan prediksi label dari set data yang diberikan.

```
[ ] # Print accuracy for training and testing
print(f'Akurasi Training: {history.history["accuracy"][-1]:.2f}')
print(f'Akurasi Testing: {history.history["val_accuracy"][-1]:.2f}')

Akurasi Training: 1.00
Akurasi Testing: 0.60
```

Gambar 16 menunjukkan hasil akurasi training dan testing. Dimana akurasi training lebih tinggi dibandingkan testing.

- a. Akurasi (*accuracy*) *presentase* dari jumlah total prediksi yang benar.

Analisis Kepuasan Pelanggan Pada Pelayanan

E-Commerce Tokopedia Dan Shopee Pada Twitter Menggunakan Recurrent Neural Network

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

$$Accuracy = \frac{4+4}{4+4+3+4} \times 100\% = \frac{8}{15} \times 100\% = 53\% / 0,53$$

b. Presisi (*precision*) adalah porsi kasus positif yang teridentifikasi dengan benar. $= 57\% / 0,57$

$$Precision = \frac{4+3}{TP + FP} \times 100\%$$

$$Precision = \frac{4}{7} \times 100\% = \frac{4}{7} \times 100\%$$

c. *Recall* merupakan *presentase* kasus positif yang teridentifikasi dengan benar.

d. *F1-Score* merupakan perbandingan rata-rata presisi dan *recall* yang dibobot.

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

$$(recall + precision) \quad TP \quad f1 score = 2 \times \frac{4}{4+4} \times 100\% = \frac{4}{8} \times 100\% \\ (recall \times precision) \quad \times 100\% \quad f1 score = 2 \times \frac{(0,5 \times 0,57)}{(0,5 + 0,57)} \times 100\% = 50\% / 0,5$$

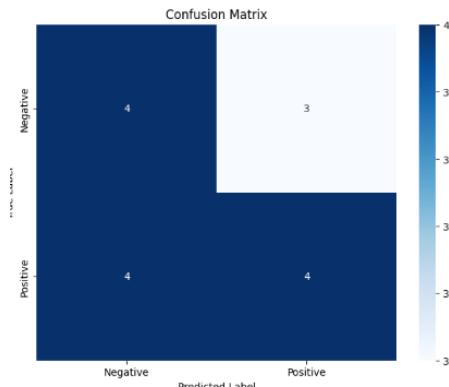
Keterangan : True Positive (TP) = 3

$$f1 score = 2 \times \frac{0,285}{1,07} \times 100\% = 57\%$$

True Negative (TN) = 4

False Positive(FP) = 6

False Negative(FN) = 2



Gambar 17 Gambar 4.16 Hasil RNN Tokopedia dalam Confusion Matrix

Confusion matrix digunakan untuk mengevaluasi kinerja sebuah model klasifikasi dengan membandingkan label asli dan prediksi label dari set data yang diberikan.

4. KESIMPULAN

Berdasarkan pembahasan dan hasil analisis, bahwa ulasan tweet yang berjumlah 200 data Tokopedia dan 200 data Shopee. Data yang dibagi menjadi dua kelas yaitu kelas negatif yang memberi opini keluhan dan kecewa dari user Twitter atas pelayanan kedua e-commerce dan kelas positif menunjukkan opini kepuasan dan kesenangan atas pelayanan kedua ecommerce. Dalam proses pengujian klasifikasi dari hasil perhitungan confusion matrix yang menjadi acuan untuk prediksi dan aktual, lalu menggunakan model algoritma yaitu Simple RNN. Pada ulasan pelanggan lebih banyak menyukai pelayanan Tokopedia dari user Twitter dapat dilihat dengan data uji yang positif sebesar 140 data sedangkan data uji yang positif pada pelayanan Shopee sebesar 120 data. Hasil akurasi yang memperoleh performa lebih baik yaitu Tokopedia menggunakan algoritma Simple RNN dengan nilai akurasi mencapai 53%, presisi mencapai 57%, recall mencapai 50%, dan f1-score mencapai 53,27%. Jika dibandingkan nilai

klasifikasi terbaik dari data Shopee dari algoritme Simple RNN menunjukkan nilai klasifikasi dengan hasil akurasi sebesar 47%, presisi mencapai 60%, recall sebesar 33%, dan f1-score mencapai 40%.

5. REFERENSI

- [1] Indraloka, D. S. dan Santosa, B. (2017) "Penerapan Text Mining untuk Melakukan Clustering Data Tweet Shopee Indonesia," *Jurnal Sains dan Seni ITS*, 6(2), hal. 6–11. doi: 10.12962/j23373520.v6i2.24419.
- [2] Malik, M. dan Rohandi, A. (2017) "Effectiveness C2C E-Commerce Media In Bandung (Case study at Tokopedia.com and Bukalapak.com)," *Jurnal Ekonomi dan Bisnis-Manajemen*, 1(3 pp), hal. 177–197.
- [3] Nugroho, A. K. dan Sari, P. K. (2016) "Analisis Pengaruh Kualitas Website Tokopedia Terhadap Kepuasan Pengguna Menggunakan Metode Webqual 4.0," *e-Proceeding of Management*, 3(Desember 2016), hal. 2930–2937.
- [4] Adam, R. (2019). Word Embedding Bahasa Indonesia Menggunakan Fasttext. Retrieved From <Https://Structilmy.Com/2019/04/Word- Embedding-Bahasa-Indonesia-Menggunakan-Fasttext-Part-1>.
- [5] Zulfa, I. dan Winarko, E. (2017) "Sentimen Analisis Tweet Berbahasa Indonesia Dengan Deep Belief Network," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 11(2), hal. 187. doi: 10.22146/ijccs.24716.
- [6] Yahya dan Mahpuz (2019)"Penggunaan Algoritma K-Mean Untuk Menganalisis Pelanggan Potensial Pada Dealer SPS Motor Honda Lombok Timur Nusa Tenggara Barat," *Jurnal Informatika dan Teknologi*, 2(2), Juli 2019, hal 109 – 118.
- [7] Dena Dwi Wijayanti, Feti Fatimah, Budi Santoso (2020) "Analisis Tingkat Kepuasan Berdasarkan E-Service Quality Pelanggan Pada Toko Online(Tokopedia)," Oktober2020.<http://repository.unmuhjember.ac.id/id/eprint/6233>
- [8] Wijayanto, I. R., Cholissodin, I. dan Sari, Y. A. (2021) "Pengaruh Metode Word Embedding dalam Vector Space Model Pada Pemerolehan Informasi Materi IPA Siswa SMP," *Jurnal Pengembangan Teknologi Informasi dan IlmuKomputer*, 5(3), hal. 950–959.
- [9] Thenata, A. P. (2021) "Text Mining Literature Review on Indonesian Social Media," *Jurnal Edukasi dan Penelitian Informatika*, 7(2), hal. 226–232.
- [10] Lupi, F. R. dan Nurdin (2016) "Analisis strategi pemasaran dan penjualan," *Jurnal Elektronik Sistem Informasi Dan Komputer (Jesik)*, 2(1), hal. 20–29.