

2. Cleaning

```
import pandas as pd
pd.set_option('display.max_columns', None)
df = pd.read_csv('dataset/shoelogo/shoelogo_data_prelabel.csv')
df.head(1000)
```

	content	score	Label
0	game nya sangat bagus,book struk permainan bec...	3	Positif
1	Chikkkkk kaaaaaannn nak ngelinggga panti...	1	Negatif
2	Gem andaw best tapi enggak kaya yang dipikn	1	Negatif
3	game nya enak nya game login	1	Negatif
4	Game Jan Cok game game anak gendeng	1	Negatif
...
1495	Kenapa setiap w mau game tapi mau dapat apa L...	1	Negatif
1496		3	Positif
1497		1	Negatif
1498	Game nya sangat seru enak	3	Positif

3. Preprocessing

```
PREPROCESSING
import pandas as pd
df = pd.read_csv('dataset/shoelogo/shoelogo_data_cleaning_preprocessing.csv')
df.head(1000)
```

	content	score	Label
0	Chikkkkk kaaaaaannn nak ngelinggga panti...	1	Negatif
1	Gem andaw best tapi enggak kaya yang dipikn	1	Negatif
2	game nya enak nya game login	1	Negatif
3	Game Jan Cok game game anak gendeng	1	Negatif
4	Makud nya ap kok nya gh kenapa napa bisa ib...	1	Negatif
...
1289	BANG GWE NYA BAGUS TAPI NGELIX TERUS EROR	3	Positif
1290	Kenapa setiap w mau game tapi mau dapat apa L...	1	Negatif
1291		3	Positif
1292		1	Negatif
1293	Game nya sangat seru enak	3	Positif

4. Case folding

```
CASE FOLDING
import re
def clean_text(text):
    text = re.sub(r'[\s]+', ' ', text)
    text = re.sub(r'[\p{L}]{4,}', lambda m: m.group().lower(), text)
    text = re.sub(r'[\p{L}]{3,}', lambda m: m.group().lower(), text)
    text = re.sub(r'[\p{L}]{2,}', lambda m: m.group().lower(), text)
    return text

df['content'] = df['content'].apply(clean_text)
df.head(1000)
```

	content	score	Label	text_clean
1	Chikkkkk kaaaaaannn nak ngelinggga panti...	1	Negatif	chikkkkk kaaaaaannn nak ngelinggga panti...
2	Gem andaw best tapi enggak kaya yang dipikn	1	Negatif	gem andaw best tapi enggak kaya yang dipikn
3	game nya enak nya game login	1	Negatif	game nya enak nya game login
4	Game Jan Cok game game anak gendeng	1	Negatif	game jan cok game game anak gendeng
5	Makud nya ap kok nya gh kenapa napa bisa ib...	1	Negatif	makud nya ap kok nya gh kenapa napa bisa ib...
...
1289	BANG GWE NYA BAGUS TAPI NGELIX TERUS EROR	3	Positif	bang gwe nya bagus tapi ngelix terus eror
1495	Kenapa setiap w mau game tapi mau dapat apa L...	1	Negatif	kenapa setiap w mau game tapi mau dapat apa L...

9. Split data

```
PEMBAGIAN DATA

# Membagi data menjadi data training dan testing dengan test_size = 0.25 dan random_state nya 0
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(data_clean['content'], data_clean['label'],
                                                  test_size = 0.25,
                                                  random_state = 0)
```

L 4 TF-IDF

```
TF-IDF

+ Code + Text

| | from sklearn.feature_extraction.text import TfidfVectorizer
| |
| | tfidf_vectorizer = TfidfVectorizer()
| | tfidf_train = tfidf_vectorizer.fit_transform(X_train)
| | tfidf_test = tfidf_vectorizer.transform(X_test)
| |
| | print(X_train.shape)
| | print(y_train.shape)
| | print(X_test.shape)
| | print(y_test.shape)

| | (1063,)
| | (1063,)
| | (263,)
| | (263,)
```

```
from sklearn.feature_extraction.text import CountVectorizer

vectorizer = CountVectorizer()
vectorizer.fit(X_train)

+ CountVectorizer
CountVectorizer()

| | X_train = vectorizer.transform(X_train)
| | X_test = vectorizer.transform(X_test)
```

L 5 Naive bayes

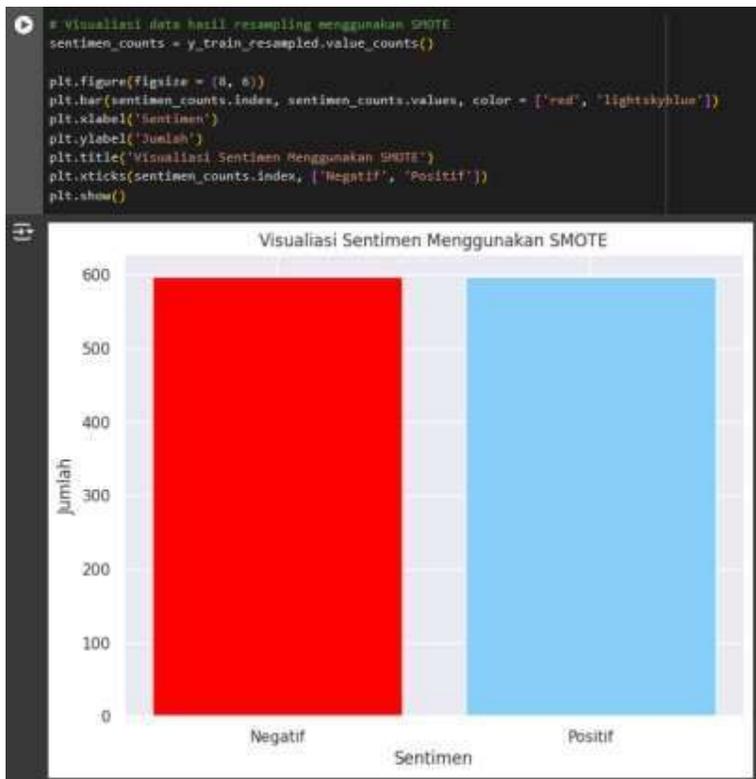
```
from sklearn.naive_bayes import MultinomialNB

nb = MultinomialNB()
nb.fit(tfidf_train, y_train)

+ MultinomialNB
MultinomialNB()
```

L 6 Smote

```
# Melakukan oversampling SMOTE pada data latih
smote = SMOTE(random_state = 0)
X_train_resampled, y_train_resampled = smote.fit_resample(X_train, y_train)
```



10. Evaluasi model

```

from sklearn.metrics import accuracy_score
accuracy = accuracy_score(y_test, y_pred)

import pandas as pd
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
from sklearn.metrics import confusion_matrix
import matplotlib.pyplot as plt

xif = MultinomialNB()
xif.fit(x_train, y_train)
predicted = xif.predict(x_test)

print("\nConfusion Matrix Accuracy:", accuracy_score(y_test, predicted))
print("\nConfusion Matrix Precision:", precision_score(y_test, predicted), average="binary", pos_label="Negatif")
print("\nConfusion Matrix Recall:", recall_score(y_test, predicted), average="binary", pos_label="Negatif")
print("\nConfusion Matrix F1 score:", f1_score(y_test, predicted), average="binary", pos_label="Negatif")

print("\nConfusion matrix in (confusion matrix, y_test, predicted):")
print("-----")
print(classification_report(y_test, predicted, zero_division=0))

# Calculate confusion matrix
cm = confusion_matrix(y_test, predicted)

# Assign confusion matrix values
tn, fp, fn, tp = cm.ravel()

print("True Positive : ", tp)
# accuracy score
print("True Negative : ", tn)
# Precision score
print("True Positive : ", fp)
# Recall score
print("True Negative : ", fn)

```

```

MultinomialNB Accuracy: 0.8620689655172413
MultinomialNB Precision: 0.8461538461538461
MultinomialNB Recall: 0.7777777777777778
MultinomialNB f1_score: 0.8105263157894737
confusion_matrix:
[[ 77  22]
 [ 14 148]]
=====

```

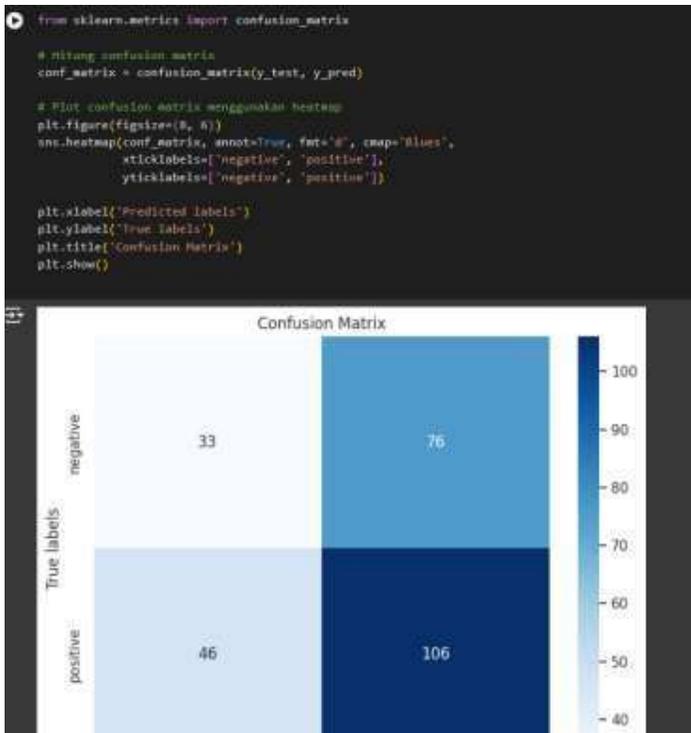
	precision	recall	f1-score	support
Negatif	0.85	0.78	0.81	99
Positif	0.87	0.91	0.89	162
accuracy			0.86	261
macro avg	0.86	0.85	0.85	261
weighted avg	0.86	0.86	0.86	261

```

True Positive : 148
True Negative : 77
False Positive : 22
False Negative : 14

```

11. Confusion matrix





Jadwal penelitian

No	Kegiatan	Februari	Maret	April	Mei	Juni	Juli
1	Pengajuan judul rta						
2	Persiapan menyusun proposal						
3	Pengumpulan data						
4	Pengolahan dan analisis data						
5	Implementasi algoritma analisis sentimen						
6	Analisis hasil						
7	Penyusunan laporan						
8	Sidang seminar hasil						

Research Results of Sentiment Analysis of "Stumble Guys" Game Reviews on Play Store



Chantal meunier · 201700447020@unil.ac.id
@support

6:55 PM (1 minutes ago) ☆ ↶ ⋮

To the Development Team of "Stumble Guys",

My name is Awanj Hejuna Nurdy, a student and researcher interested in sentiment analysis of user reviews on the Google Play Store. I have conducted a study titled "**Sentiment Analysis of Stumble Guys Game Reviews on Play Store Using Naive Bayes Algorithm**". In this research, I used the Naive Bayes algorithm to classify user reviews into positive and negative sentiments.

Here is a summary of my research findings:

- **Accuracy:** 86.21%
- **Positive Precision:** 87%
- **Negative Precision:** 35%
- **Positive Recall:** 91%
- **Negative Recall:** 78%
- **Positive F1-Score:** 89%
- **Negative F1-Score:** 81%

With a fairly high accuracy rate, this model is able to accurately predict the majority of reviews. These results can provide useful insights into how users feel and rate the game "Stumble Guys".

I hope that the findings of this research can assist the development team in improving the game's quality and user experience.

Thank you for your attention and cooperation.

One attachment · Scanned by Gmail



KARTU KENDALI BIMBINGAN LAPORAN KARYA ILMIAH

Nama Mahasiswa : Awang herjunie nurdy
 NIM : 2011102441020
 Nama Dosen Pembimbing : Abdul Rahim, S.Kom., M.Cs
 Judul Penelitian : ANALISIS SENTIMEN ULASAN GAME STUMBLE GUYS
 PADA PLAYSTORE MENGGUNAKAN ALGORITMA NAÏVE
 : BAYES

No	Tanggal	Uraian Pembimbing	Paraf Dosen
1	May 2, 2024	bimbingan bab 1 dan menambahkan jurnal international pada latar belakang	
2	May 7, 2024	revisi rumusan masalah dan manfaat penelitian	
3	May 13, 2024	revisi penulisan dan sesuaikan dengan pedoman penulisan	
4	May 16, 2024	bimbingan bab 2 alur penelitian dan obyek penelitian	
5	May 23, 2024	revisi tahapan penelitian yang masih kurang	
6	June 3, 2024	revisi pada obyek penelitian, tentang stumble guys	
7	June 5, 2024	jadwal penelitian harus menyesuaikan waktu dengan awal mulai semester dan menambahkan jadwalnya sesuai dengan kegiatan	

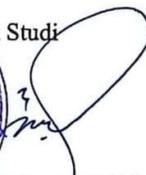
8	June 10, 2024	bimbingan bab 3 dan revisi pada preprocessing menambahkan hasil preprocessing	
9	June 24, 2024	bimbingan bab 3, dan memperbaiki format penulisan menghapus point-point menggunakan (i),(ii) dan (iii) untuk sub bab	
10	June 25, 2024	bimbingan bab 3 dan 4 simpulan dan saran dan merapikan penulisan dengan rata kiri dan paragraf	

Dosen Pembimbing



Abdul Rahim, S.Kom., M.Cs
(NIDN.0009047901)

Mengetahui

Abdul Rahim, S.Kom., M.T.I.
(NIDN.1118019202)

SKRIPSI AWANG HERJUNIE NURDY

by Teknik Informatika Universitas Muhammadiyah Kalimantan Timur

Submission date: 17-Jul-2024 01:37PM (UTC+0800)

Submission ID: 2418113338

File name: SKRIPSI_AWANG_HERJUNIE_NURDY_-_AWANG_HERJUNIE_NURDY.docx (1.42M)

Word count: 3546

Character count: 22584

SKRIPSI AWANG HERJUNIE NURDY

ORIGINALITY REPORT

23%

SIMILARITY INDEX

19%

INTERNET SOURCES

14%

PUBLICATIONS

6%

STUDENT PAPERS

PRIMARY SOURCES

1	journal.ipm2kpe.or.id Internet Source	3%
2	ejurnal.stie-trianandra.ac.id Internet Source	2%
3	Imam Fathur Rahman, Anisa Nur Hasanah, Nono Heryana. "ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI SAMSAT DIGIITAL NASIONAL (SIGNAL) DENGAN MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER", Jurnal Informatika dan Teknik Elektro Terapan, 2024 Publication	2%
4	repository.uin-suska.ac.id Internet Source	2%
5	repository.upi.edu Internet Source	1%
6	dspace.umkt.ac.id Internet Source	1%
7	ejurnal.stmik-budidarma.ac.id Internet Source	1%

RIWAYAT HIDUP



Nama lengkap penulis adalah awang herjunie nurdy, Lahir di kota tenggarong, 24 juni 2002. Penulis merupakan anak ke 1 dari 3 bersaudara, penulis berkebangsaan indonesia dan suku kutai dan beragama islam, saat ini penulis tinggal di tenggarong, kalimantan timur. Penulis memulai pendidikan dasar di sekolah dasar sdn 004 tenggarong pada tahun 2008 sampai 2014, kemudian penulis melanjutkan pendidikan di sekolah menengah pertama negeri 2 tenggarong pada tahun 2014 hingga tahun 2017, dan setelah itu melanjutkan sekolah menengah kejuruan negeri 2 tenggarong pada tahun 2018 hingga tahun 2020 pendidikan dilanjutkan di universitas muhammadiyah kalimantan timur dan mengambil program studi S1 Teknik Informatika pada tahun 2020.