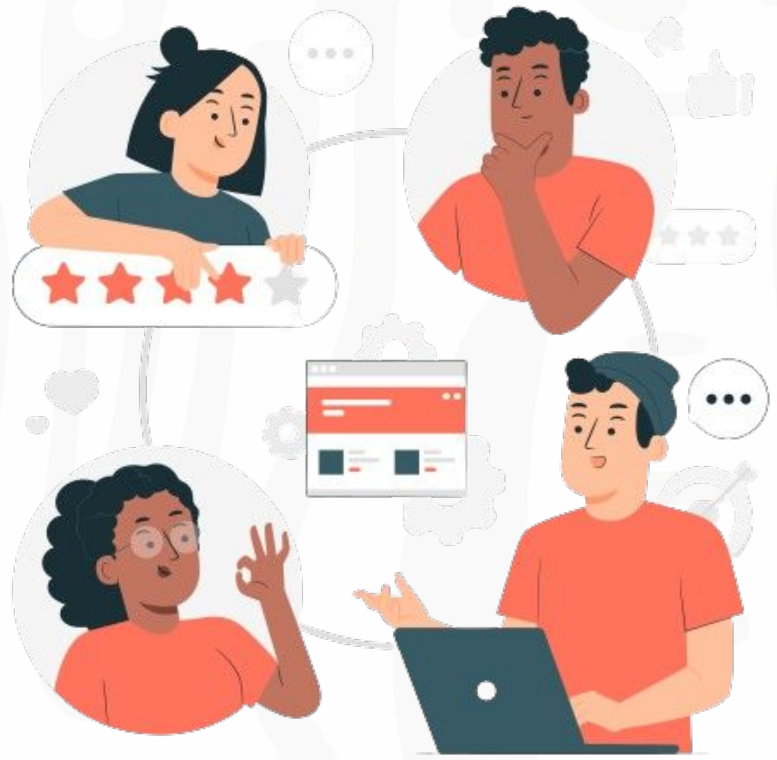
Sentiment Analysis Bukalapak Apps on Google Play Store Reviews

Fakhrizal Ahadiat

Table of Content



Business Understanding

Background

2022

56%
Leading the Penetration
Digital Warung Indonesia

2023

2 million
Transactions/Day

130 Million
Users

16,8 Million
MSME Partners

Bukalapak

Background

"Every e-commerce has its advantages and disadvantages. Usually, the experience of the users of the platform is always expressed in the comment section, be it criticism or satisfaction. This factor usually plays an important role whether the platform is good or not and it can affect the brand growth and quality of the e-commerce platform"

(Zhu & Zang, 2010) The benefits of online reviews as a good medium for disseminating information and have been shown to influence customer purchasing decisions

Business Understanding

Problem Statement

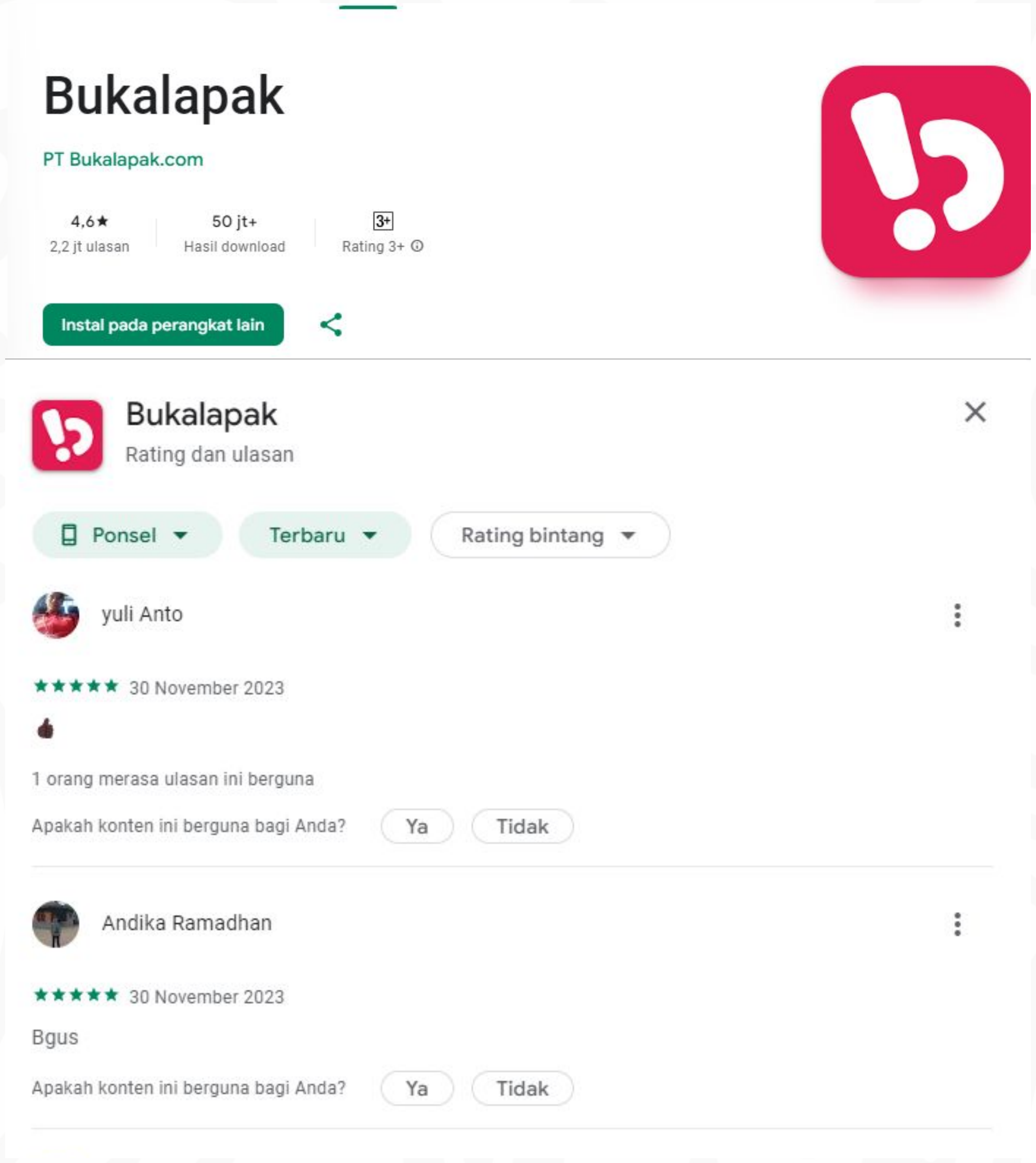
- How customers evaluate Bukalapak application whether it includes positive or negative sentiments
- Knowing the problems that exist on the platform based on negative user reviews
- What strategies can be done by the Bukalapak platform to improve the performance of the application

Objective

- Knowing the problems faced based on negative user reviews
- Create the best classification model that can predict user sentiment based on existing reviews so that the platform can immediately anticipate and make improvements.
- Provide recommendations to improve the quality and performance of the platform

Data Gathering

Data Gathering



Dataset obtained from scraping on Google Play Store website 28 November 2023

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30000 entries, 0 to 29999
Data columns (total 11 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   reviewId              30000 non-null  object
 1   userName              30000 non-null  object
 2   userImage             30000 non-null  object
 3   content               30000 non-null  object
 4   score                 30000 non-null  int64
 5   thumbsUpCount         30000 non-null  int64
 6   reviewCreatedVersion  22043 non-null  object
 7   at                    30000 non-null  object
 8   replyContent          10671 non-null  object
 9   repliedAt             10671 non-null  object
 10  appVersion            22043 non-null  object
dtypes: int64(2), object(9)
memory usage: 2.5+ MB
```

Total data from scraping 30000 data

Data Understanding & Preprocessing

Data Understanding & Preprocessing

Dataset Information

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30000 entries, 0 to 29999
Data columns (total 11 columns):
 #   Column                Non-Null Count  Dtype
---  ------                --------------  -----
 0   reviewId              30000 non-null  object
 1   userName              30000 non-null  object
 2   userImage             30000 non-null  object
 3   content               30000 non-null  object
 4   score                 30000 non-null  int64
 5   thumbsUpCount         30000 non-null  int64
 6   reviewCreatedVersion  22043 non-null  object
 7   at                    30000 non-null  object
 8   replyContent          10671 non-null  object
 9   repliedAt             10671 non-null  object
 10  appVersion            22043 non-null  object
dtypes: int64(2), object(9)
memory usage: 2.5+ MB
```

Missing Value

Variables	Types	Count	Uniques	Missing	Percent
reviewId	object	30000	28756	0	0.000000
userName	object	30000	17544	0	0.000000
userImage	object	30000	17776	0	0.000000
content	object	30000	28701	0	0.000000
score	int64	30000	5	0	0.000000
thumbsUpCount	int64	30000	258	0	0.000000
reviewCreatedVersion	object	30000	323	7957	26.523333
at	object	30000	28746	0	0.000000
replyContent	object	30000	3607	19329	64.430000
repliedAt	object	30000	10166	19329	64.430000
appVersion	object	30000	323	7957	26.523333

Duplicated Data

```
review.duplicated(subset='reviewId').sum()
```

```
1244
```

- **The dataset consists of 11 columns and 30000 rows**
- **There are missing values in 4 columns with a total number of 27,286**
- **There are 1244 duplicate data**

Data Description:

- reviewID : User ID review
- userName : name of user who reviewed
- userImage : URL image reviewer
- content : text review
- score : rating score
- thumbsUpCount : supported and preferred reviews
- reviewCreatedVersion : review is based on the application version
- at : date review was made
- replyContent : review reply from provider
- repliedAt : date review reply was made
- appVersion : version of apps

Data Understanding & Preprocessing

```
reviewId     0
userName     0
content      0
score        0
at           0
dtype: int64
```

	user_name	text_reviews	rating_score	create_at
	E MUAMAR S	Kenapa ribet sekali ingin mengganti alamat sur...	1	2023-11-18
	Maulana Malik Ibrahim	Tolong di perbaiki sistemnya,dari segi pengamb...	5	2023-11-20
	Muhammad Farid	Salah satu Aplikasi e-commerce yang keren untu...	5	2023-11-07

Important Feature for Analysis:

userName, Content, score, at

Data Cleaning

Drop Feature:

replyContent, repliedAt, userImage, reviewCreatedVersion, appVersion, thumbsUpCount, reviewId

Drop Duplicated Data

Convert Type Data & Renaming Column

Convert Type Data :

Column : **at**

Renaming Feature:

userName : user_name

content : text_reviews

score : rating_score

at : create_at

Text Preprocessing

Case Folding

Converts all letters in the document to lowercase. Only letters 'a' to 'z' are accepted. Characters other than letters will be removed

Tokenizing

Decompose each word or phrase into smaller words

Replace Slang, Replace Word Elongation

Because most people use slang/jargon and excessive affixes in chat/text. So, these words must be converted into normal words. The library used is indoNLP

Stopwords Filtering

Stopwords are used to remove words that are not important (not meaningful). The library that will be used is Sastrawi and stopwords from nltk.

Stemming

Stemming is used to convert words into root words. Stemming using the Sastrawi library because the reviews in the dataset use Indonesian/Bahasa.

Result of Text Preprocessing

	user_name	text_reviews	rating_score	create_at	text_clean	text_token	text_stopwords	stemmed_text
	E MUAMAR S	Kenapa ribet sekali ingin mengganti alamat sur...	1	2023-11-18	kenapa ribet sekali ingin mengganti alamat sur...	['kenapa', 'ribet', 'sekali', 'ingin', 'mengga...	['ribet', 'mengganti', 'alamat', 'surel', 'akt...	ribet ganti alamat surel aktif buka bantu suru...
	Maulana Malik Ibrahim	Tolong di perbaiki sistemnya,dari segi pengamb...	5	2023-11-20	tolong di perbaiki sistemnyadari segi pengambi...	['tolong', 'di', 'perbaiki', 'sistemnyadari', ...	['perbaiki', 'sistemnyadari', 'segi', 'pengamb...	baik sistemnyadari segi ambil presentase lapak...
	Muhammad Farid	Salah satu Aplikasi e-commerce yang keren untu...	5	2023-11-07	salah satu aplikasi ecommerce yang keren untuk...	['salah', 'satu', 'aplikasi', 'ecommerce', 'ya...	['salah', 'aplikasi', 'ecommerce', 'keren', 'k...	salah aplikasi ecommerce keren kalang oke oke ...

Created each Feature from the results of Text Preprocessing to facilitate checking

Next, remove Frequent words, less words that have no meaning because they can affect the results of the prediction.

	Word	Count
23	beli	8102
196	barang	8100
547	mudah	6716
1055	latih	6598
34	lapak	6351
5	buka	6275
1077	prakerja	5453
159	belanja	5451
6	bantu	4830
27	bayar	4102

Word	Count
greenstore	1
dulukecewa	1
repos	1
testimony	1
penjelasn	1
...	...
supay	1
responbagaimana	1
dibukasedang	1
rumahanumkmmateri	1
dibutuhkanterima	1

Sentiment Labeling

user_name	text_reviews	rating_score	create_at	text_clean	text_token	text_stopwords	stemmed_text	sentiment
E MUAMAR S	Kenapa ribet sekali ingin mengganti alamat sur...	1	2023-11-18	kenapa ribet sekali ingin mengganti alamat sur...	['kenapa', 'ribet', 'sekali', 'ingin', 'mengga...	['ribet', 'mengganti', 'alamat', 'surel', 'akt...	ribet ganti alamat aktif buka bantu suruh foto...	Negative
Maulana Malik Ibrahim	Tolong di perbaiki sistemnya,dari segi pengamb...	5	2023-11-20	tolong di perbaiki sistemnyadari segi pengambi...	['tolong', 'di', 'perbaiki', 'sistemnyadari', ...	['perbaiki', 'sistemnyadari', 'segi', 'pengamb...	baik segi ambil lapak daftar super otomatis ka...	Positive
Muhammad Farid	Salah satu Aplikasi e-commerce yang keren untu...	5	2023-11-07	salah satu aplikasi ecommerce yang keren untuk...	['salah', 'satu', 'aplikasi', 'ecommerce', 'ya...	['salah', 'aplikasi', 'ecommerce', 'keren', 'k...	salah ecommerce keren oke oke pokok accapprove...	Positive

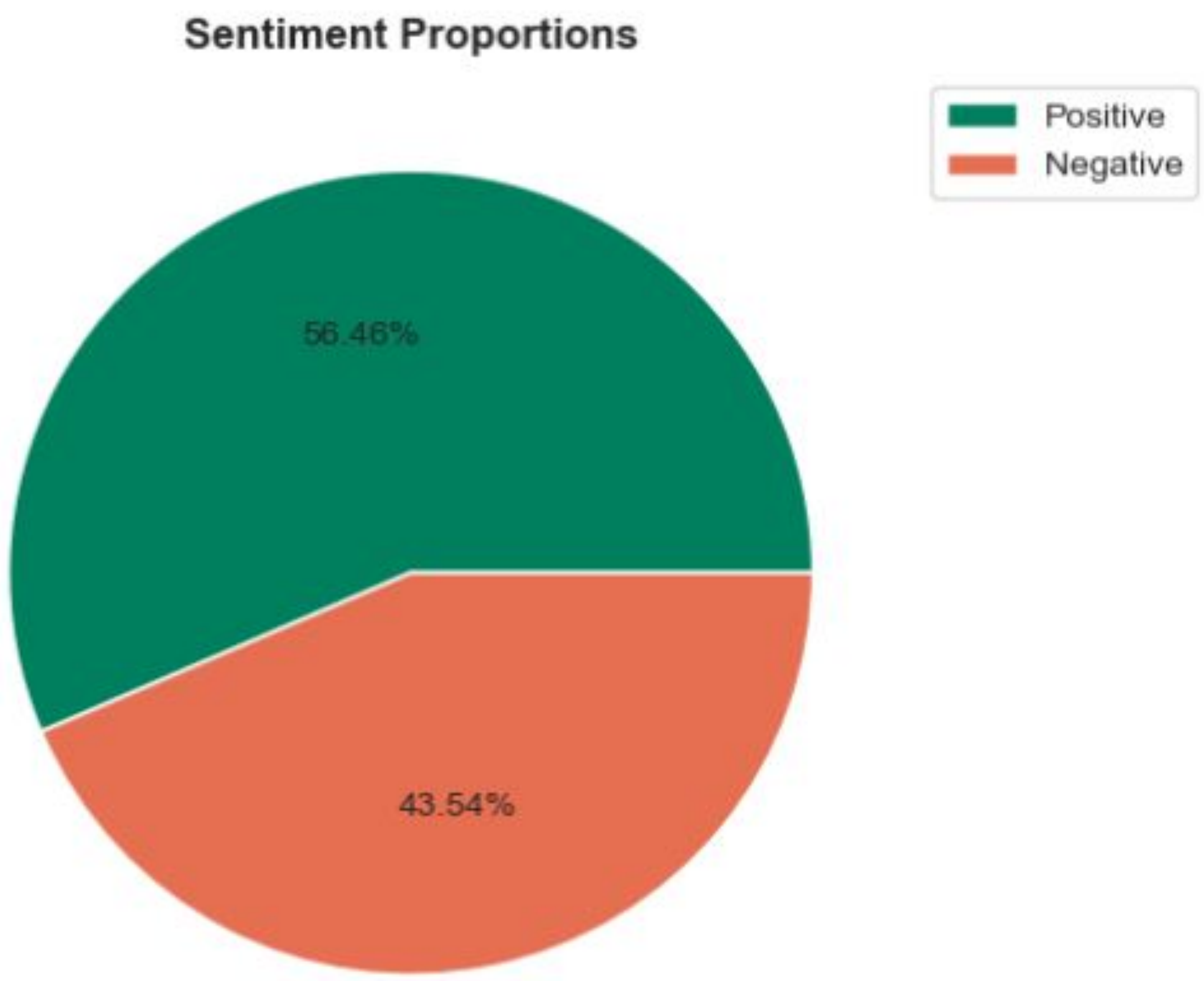
Sentiment Labeling Based on Rating

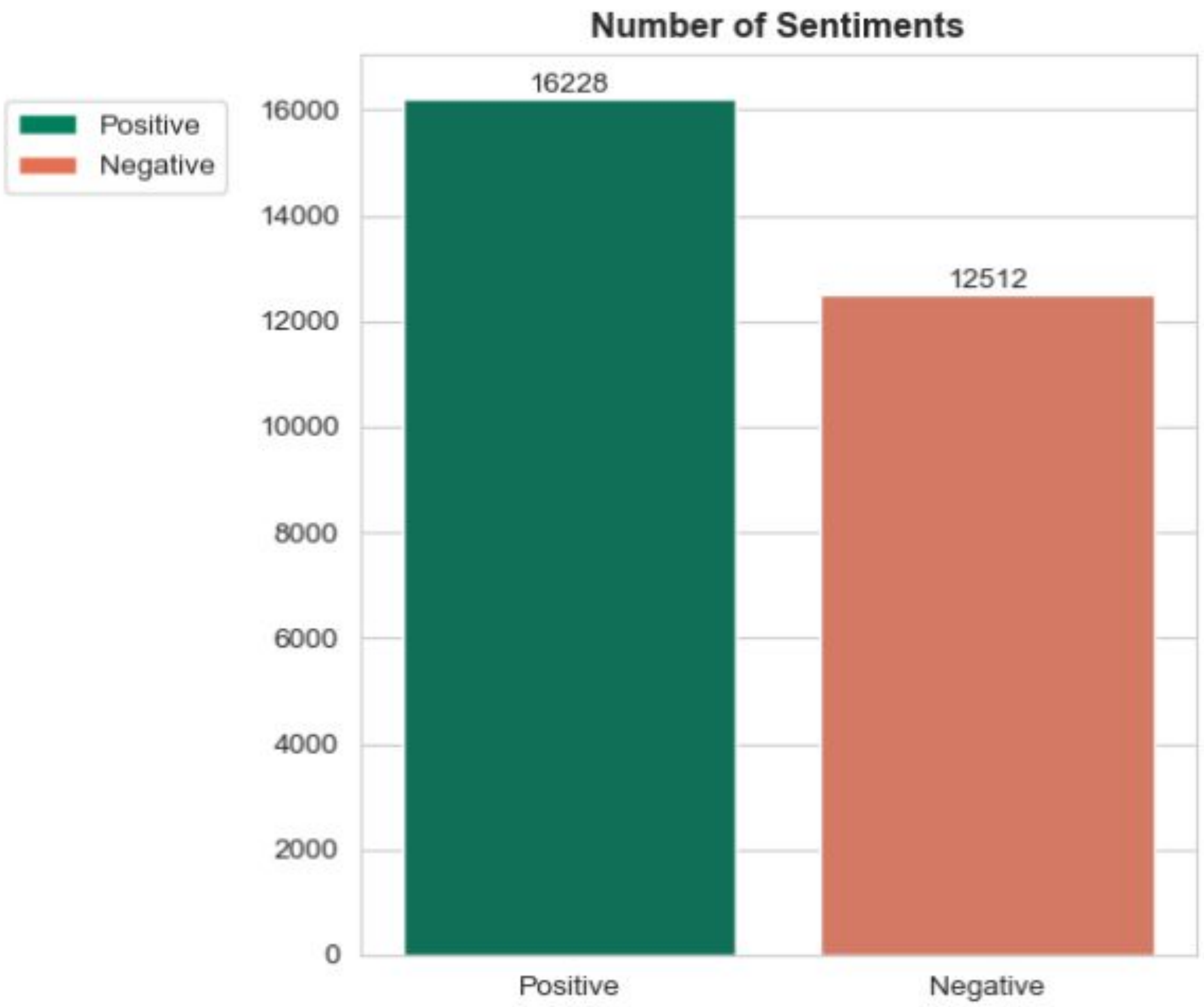
Positive : Rating 4 - 5

Negative : Rating 1 - 3

Exploratory Data Analysis

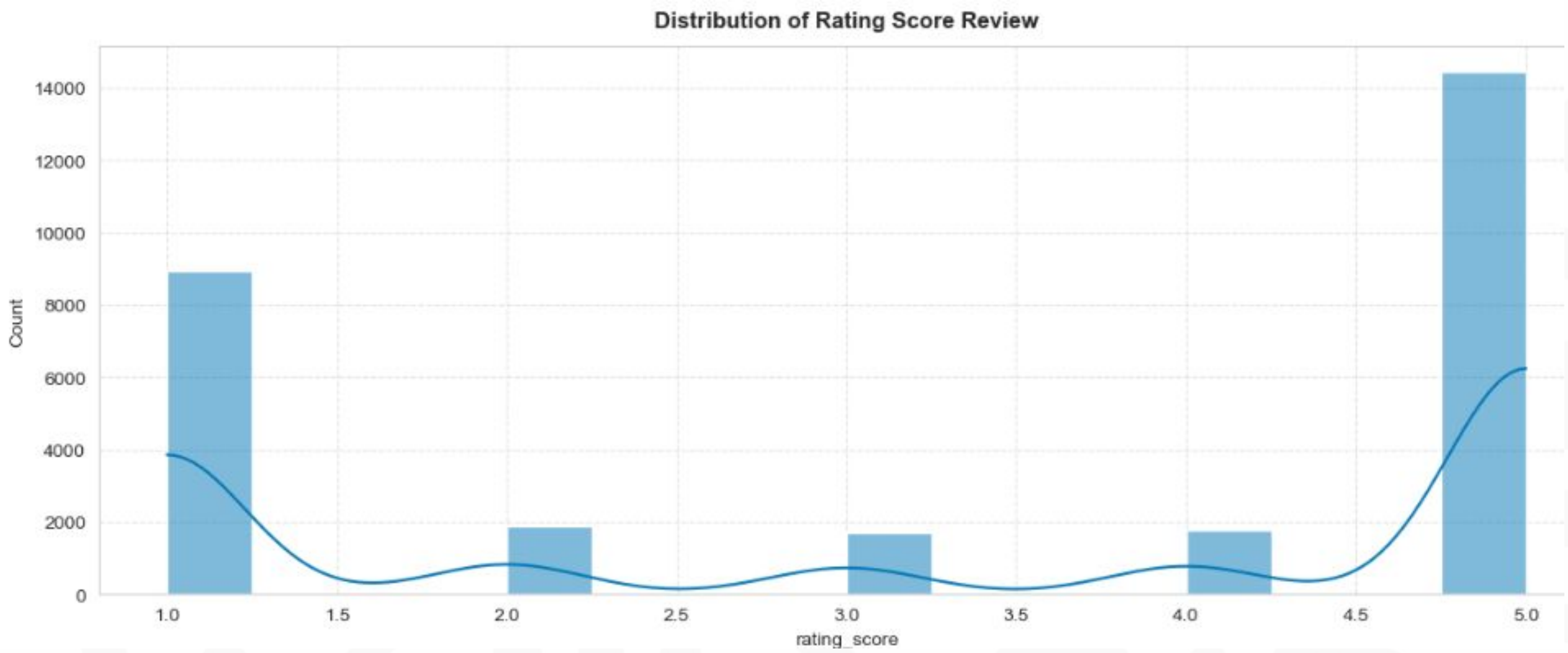
Positive Vs Negative Sentimen



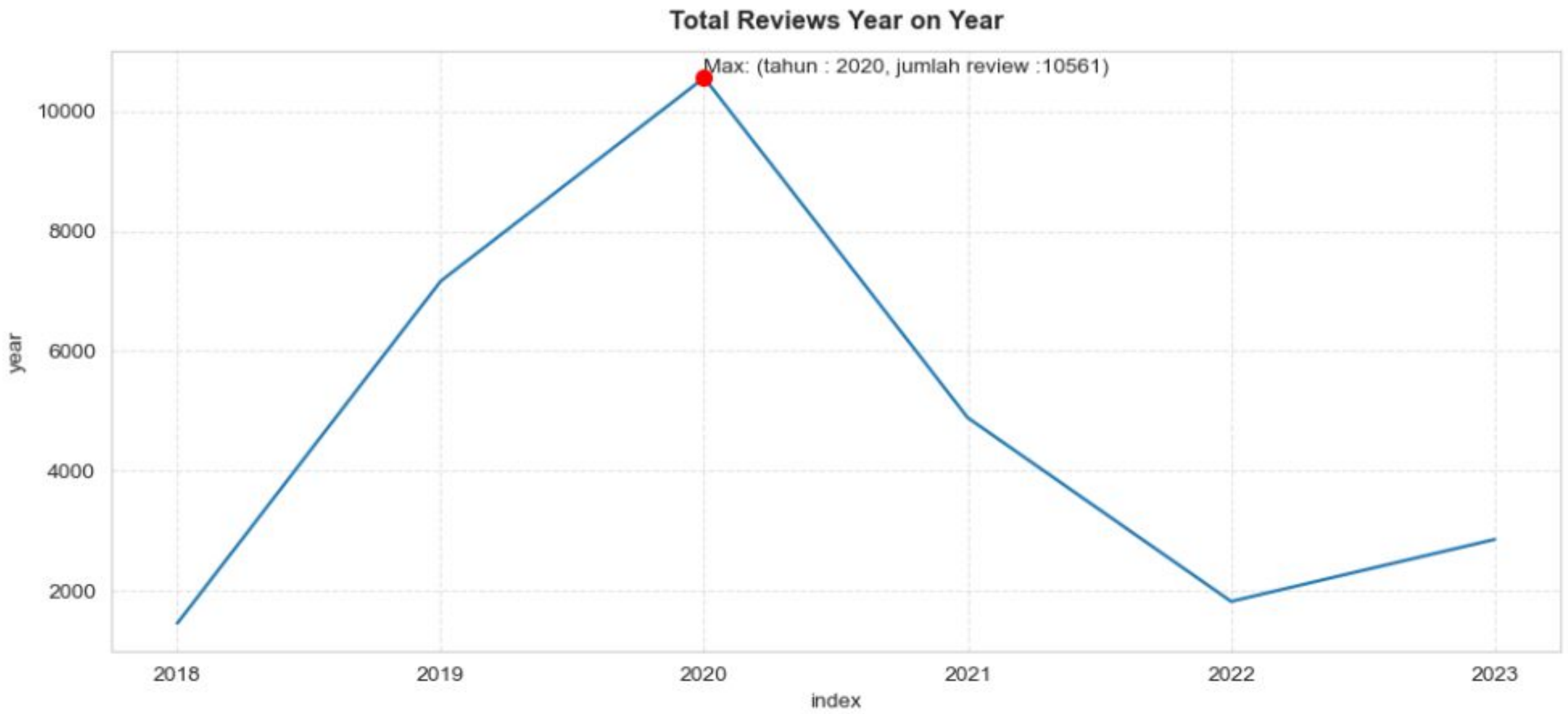


Positive sentiment is slightly more than negative sentiment. This means that some users are satisfied with the Bukalapak platform but the difference is quite small with Negative Sentiment.

Exploratory Data Analyst (EDA)



Rating By User



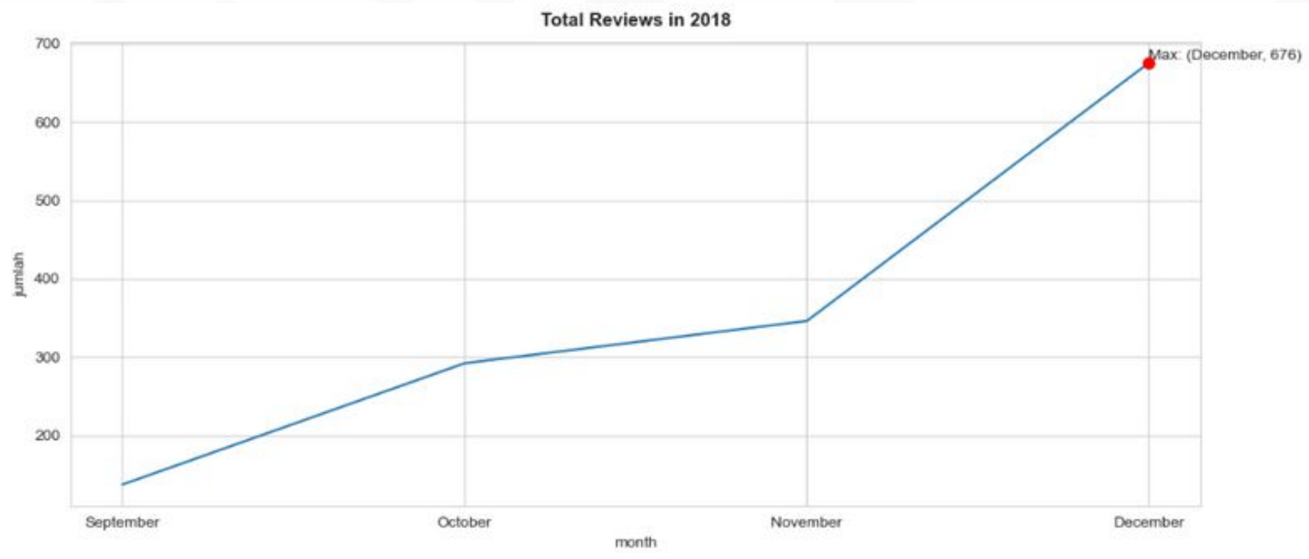
Review All Years

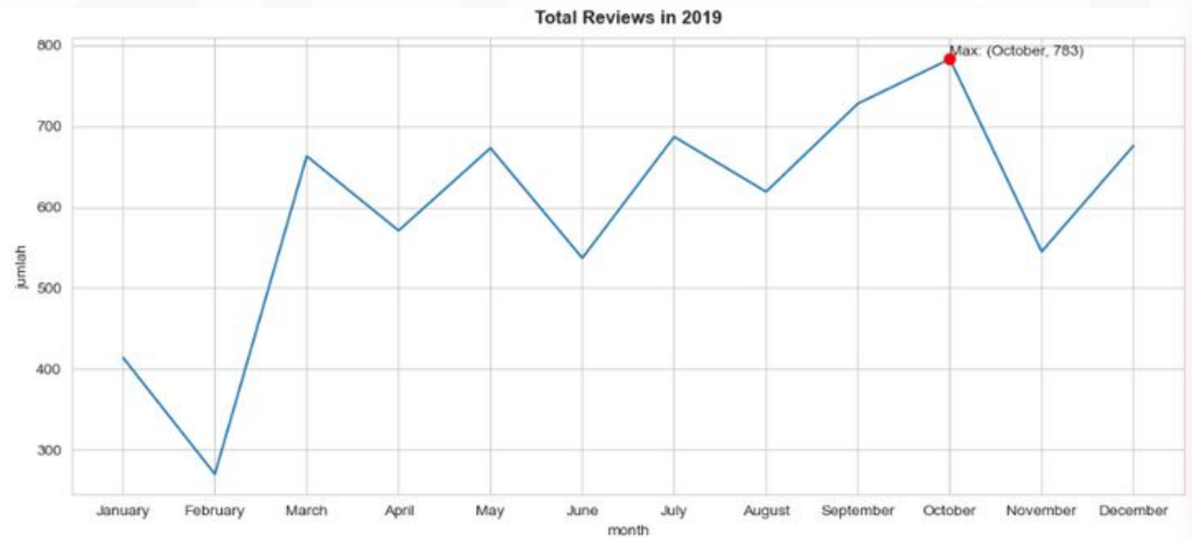
Rating distribution is mostly in 5-star and then at 1-star.

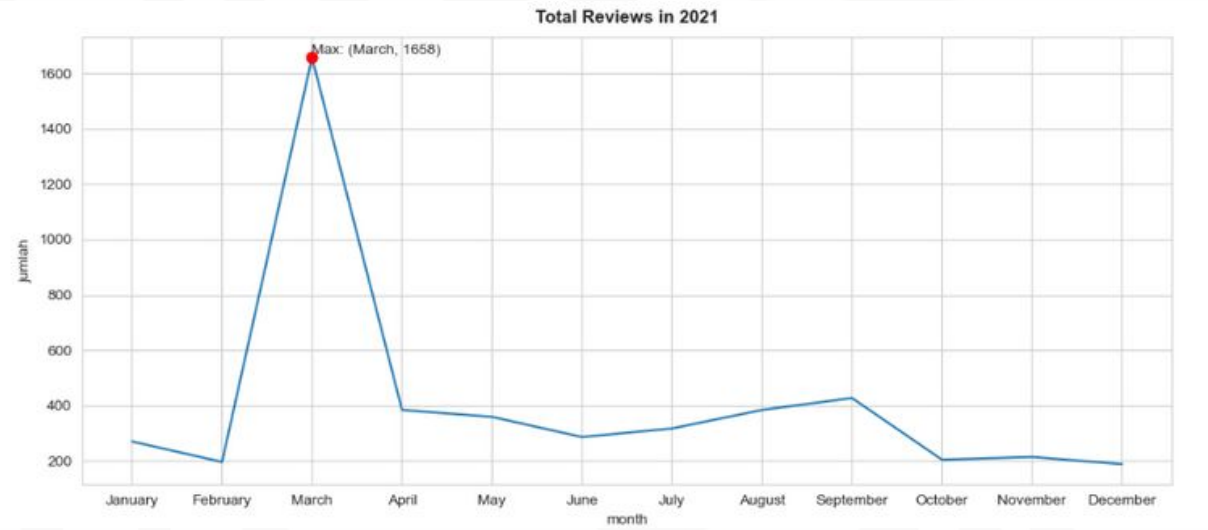
Highest total reviews in 2020

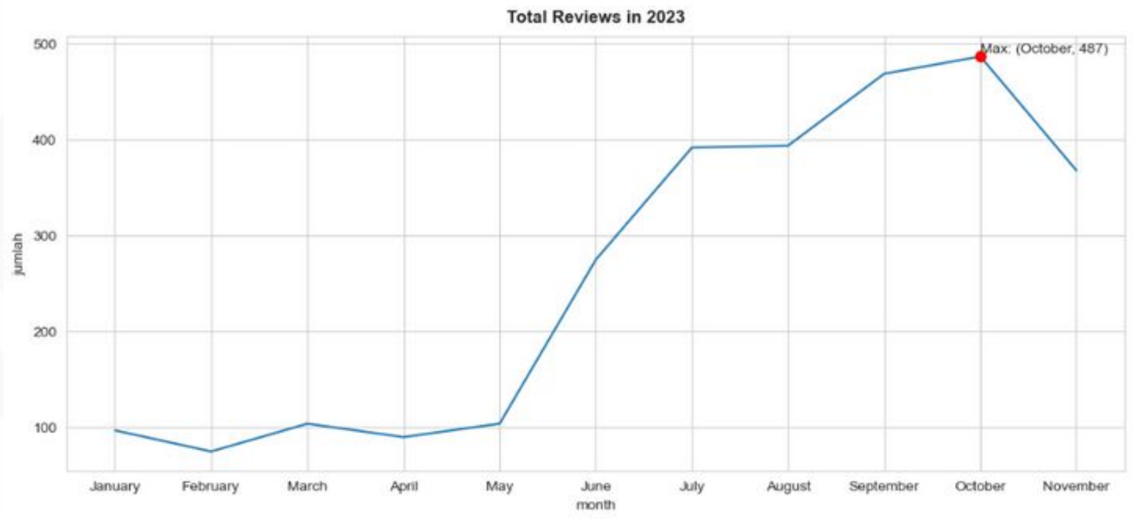
Exploratory Data Analyst (EDA)

Year to Year Reviews





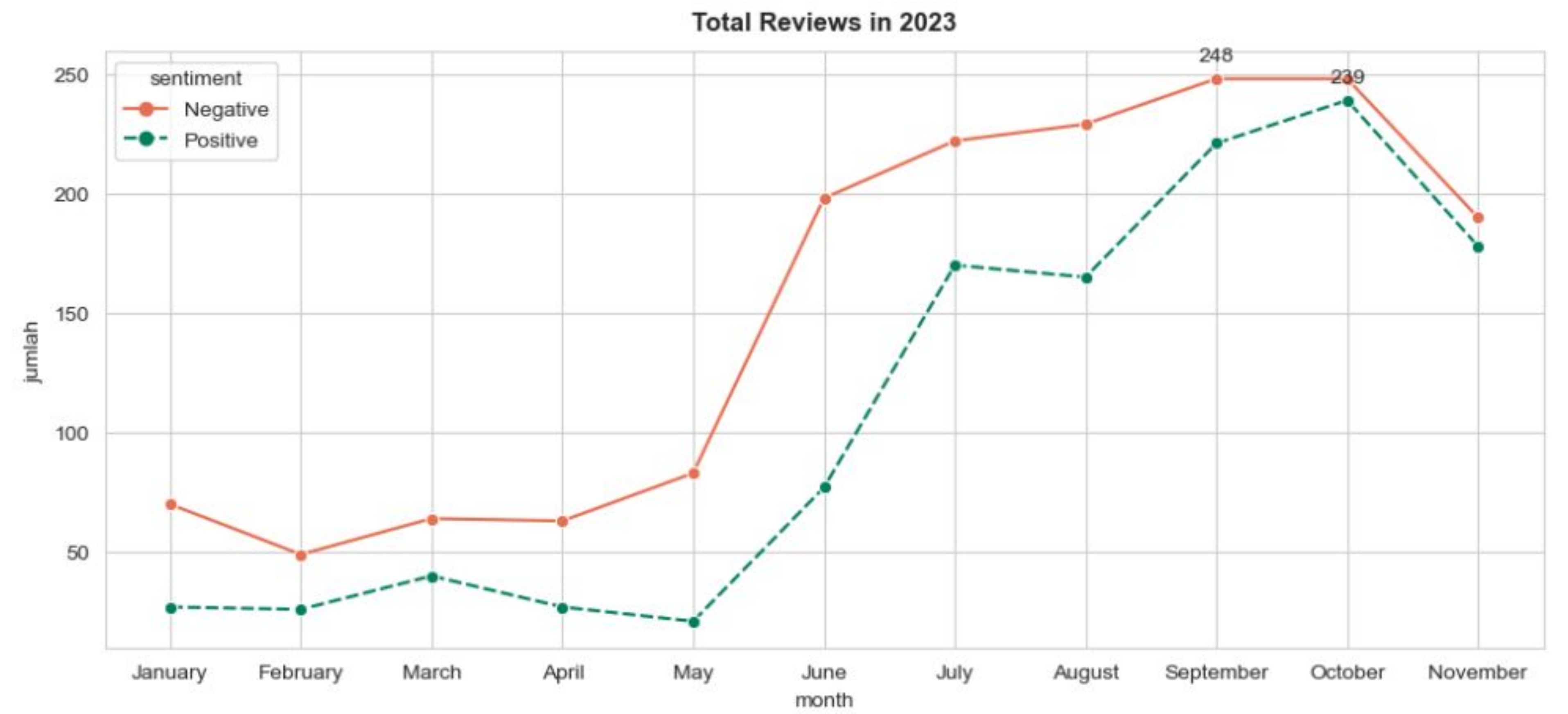




Reviews fluctuate every year

Highest total reviews in September 2020 with 4050 reviews

Exploratory Data Analyst (EDA)

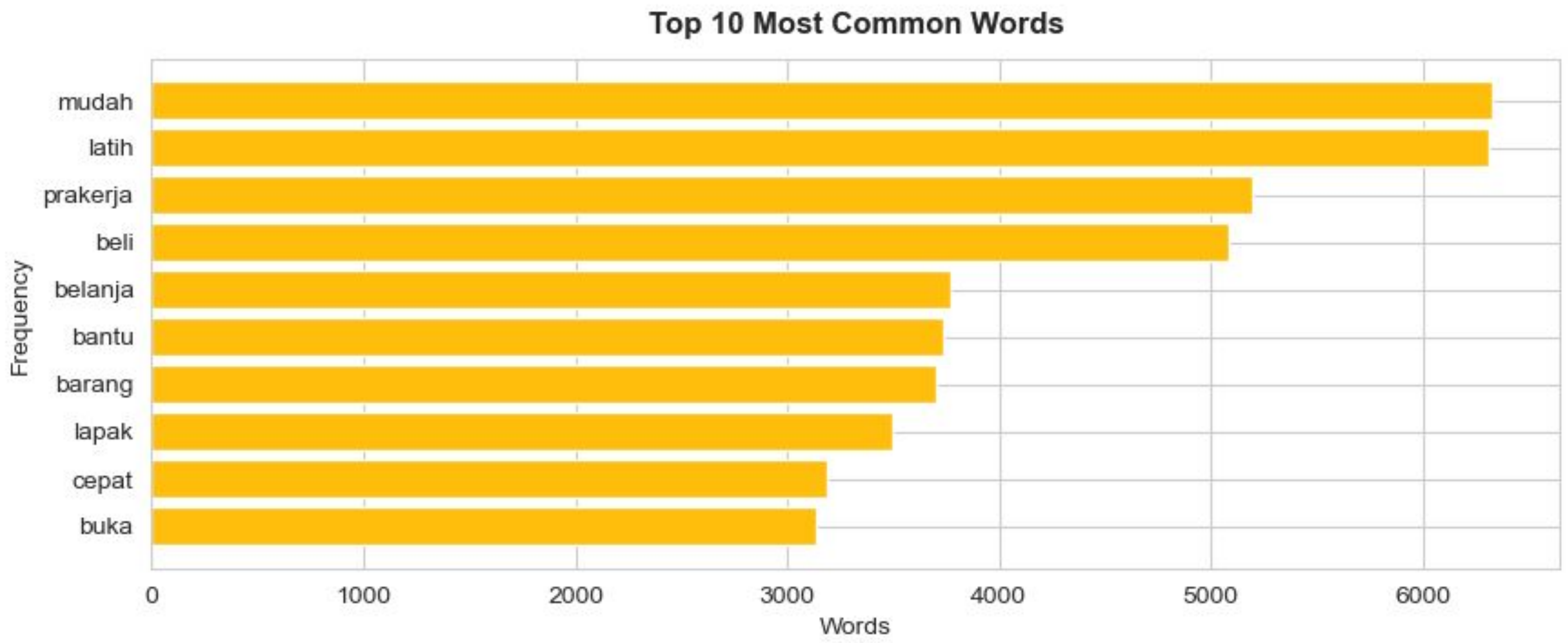


Reviews in 2023

Negative sentiment is higher than positive sentiment

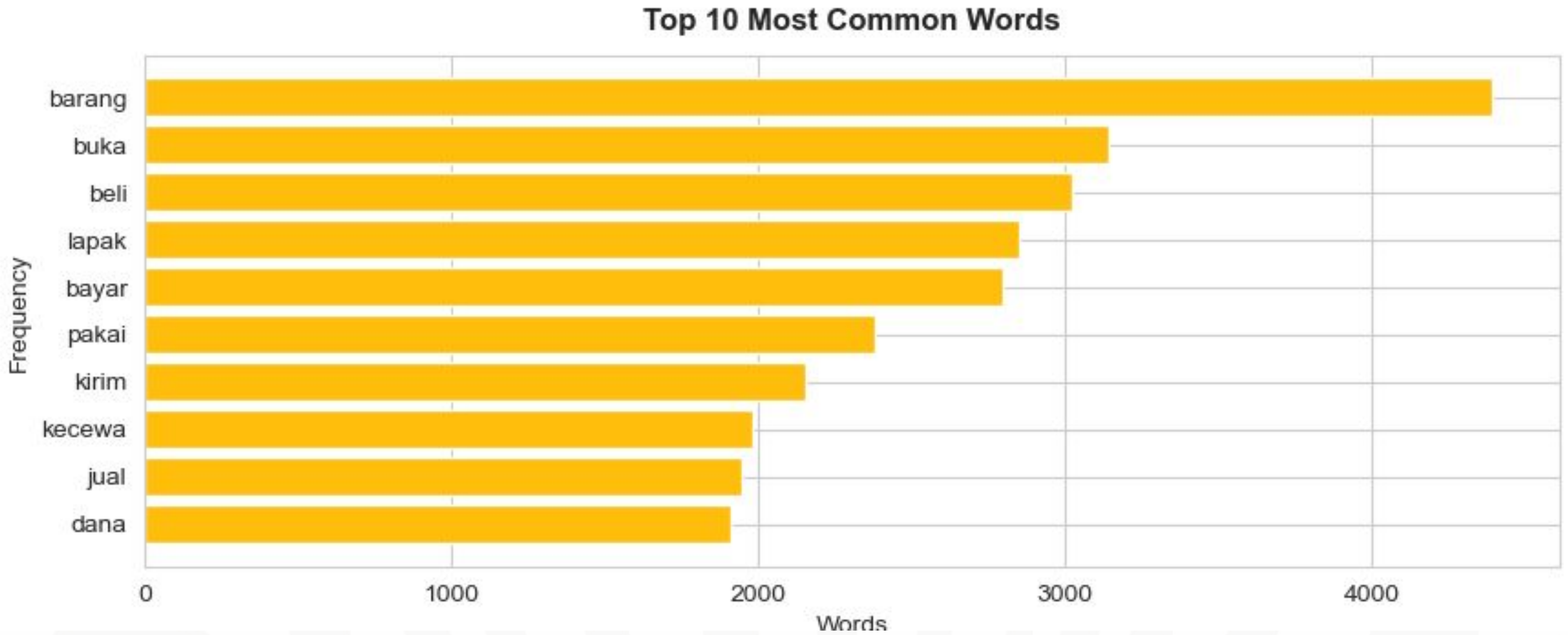
“Highest Ratio Between Positive and Negative Sentiment in October and lowest in May”

Exploratory Data Analyst (EDA)



Most Common Words on Positive Sentiment:

"terima kasih, mudah, paham, bantu, latih, prakerja"



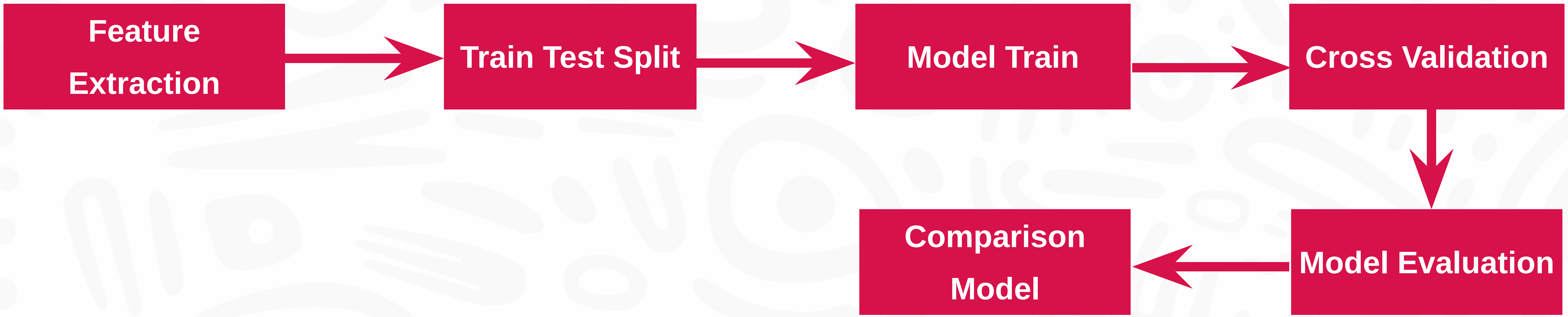
Most Common Words on Negative Sentiment:

"kecewa, bayar, pakai, komplain, barang, buka, beli"

Modeling and Evaluation

Modeling and Evaluation

Modeling Process



Feature Extraction : TFIDF Vectorizer

Train Test Data Split Ratio

Training Data : 80%

Test Data : 20%

Feature : stemmed_text

Target : label (sentiment)

Data Imbalance :

Positive : 16226

Negative : 12512

Scoring Metric : F1 Score

Modeling and Evaluation

Training Evaluation

Classifier	F1 Score	ROC AUC
LogisticRegression	0.852200	0.935444
KNeighborsClassifier	0.811800	0.877027
MultinomialNB	0.824533	0.928266
DecisionTreeClassifier	0.807300	0.784658
RandomForestClassifier	0.814880	0.928250
LGBMClassifier	0.819000	0.930862
XGBClassifier	0.821943	0.929543

Test Evaluation

Classifier	F1 Score	ROC AUC
LogisticRegression	0.862964	0.941951
KNeighborsClassifier	0.782784	0.881016
MultinomialNB	0.864439	0.934785
DecisionTreeClassifier	0.775510	0.798977
RandomForestClassifier	0.855560	0.935514
LGBMClassifier	0.854823	0.938756
XGBClassifier	0.857849	0.939223

Best 3 Model :

Logistic Regression

MultinomialNB

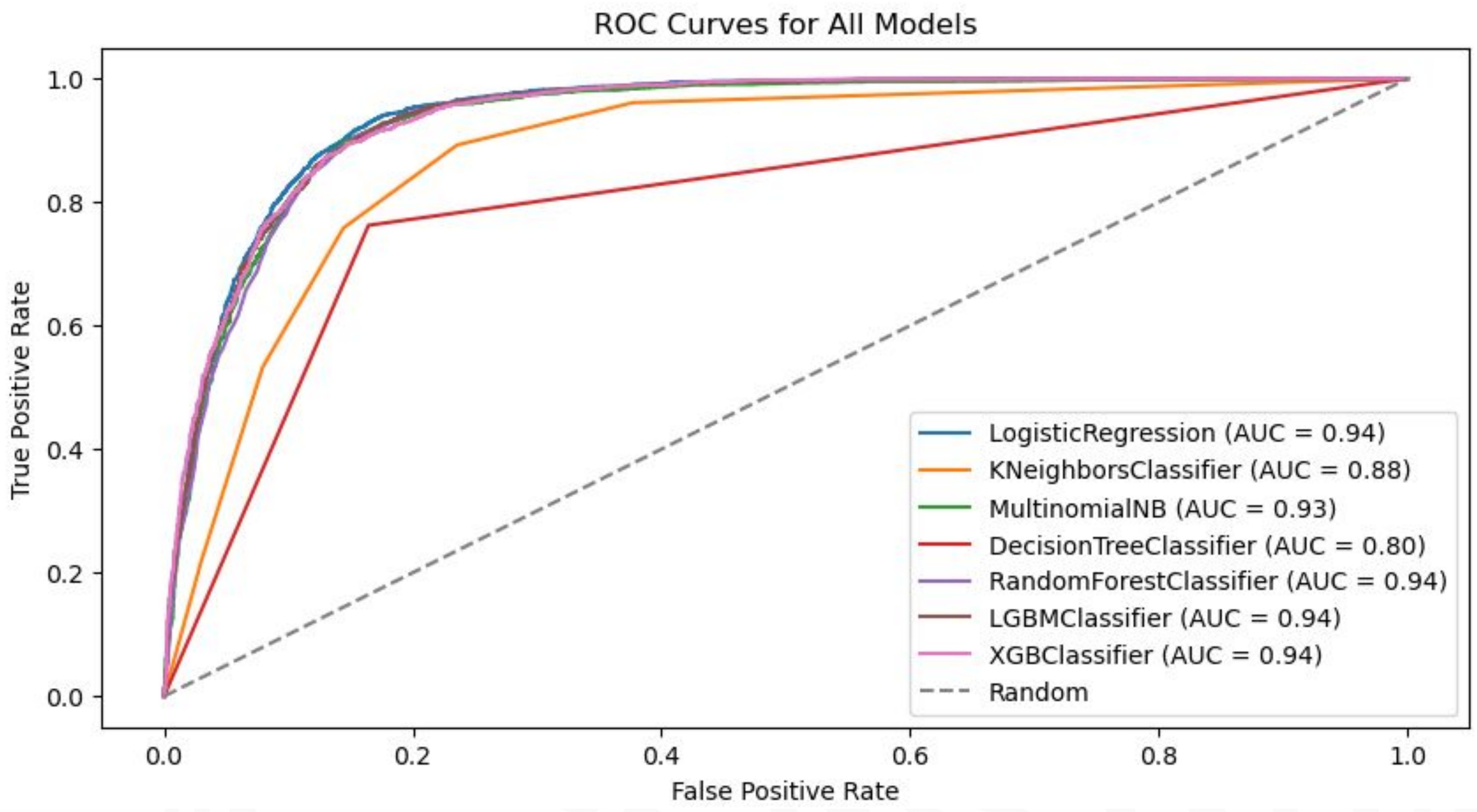
XGB Classifier

Evaluation Training using Cross Validation (Stratified K-Fold)

Model after Evaluation with Test data

Overfitting : KNN, Decision Tree

Modeling and Evaluation



Best 3 Model : Logistic Regression, MultinomialNB, XGB Classifier

ROC Threshold Score > 0.9

Modeling and Evaluation (Hyperparameter Tuning)

Hyperparameter : GridsearchCV

	Parameter	Range
TFDIF Vectorizer		
N-gram Range	ngram_range	(1,1),(1,2)
Maximum Feature	max_feature	[None, 500, 1000]

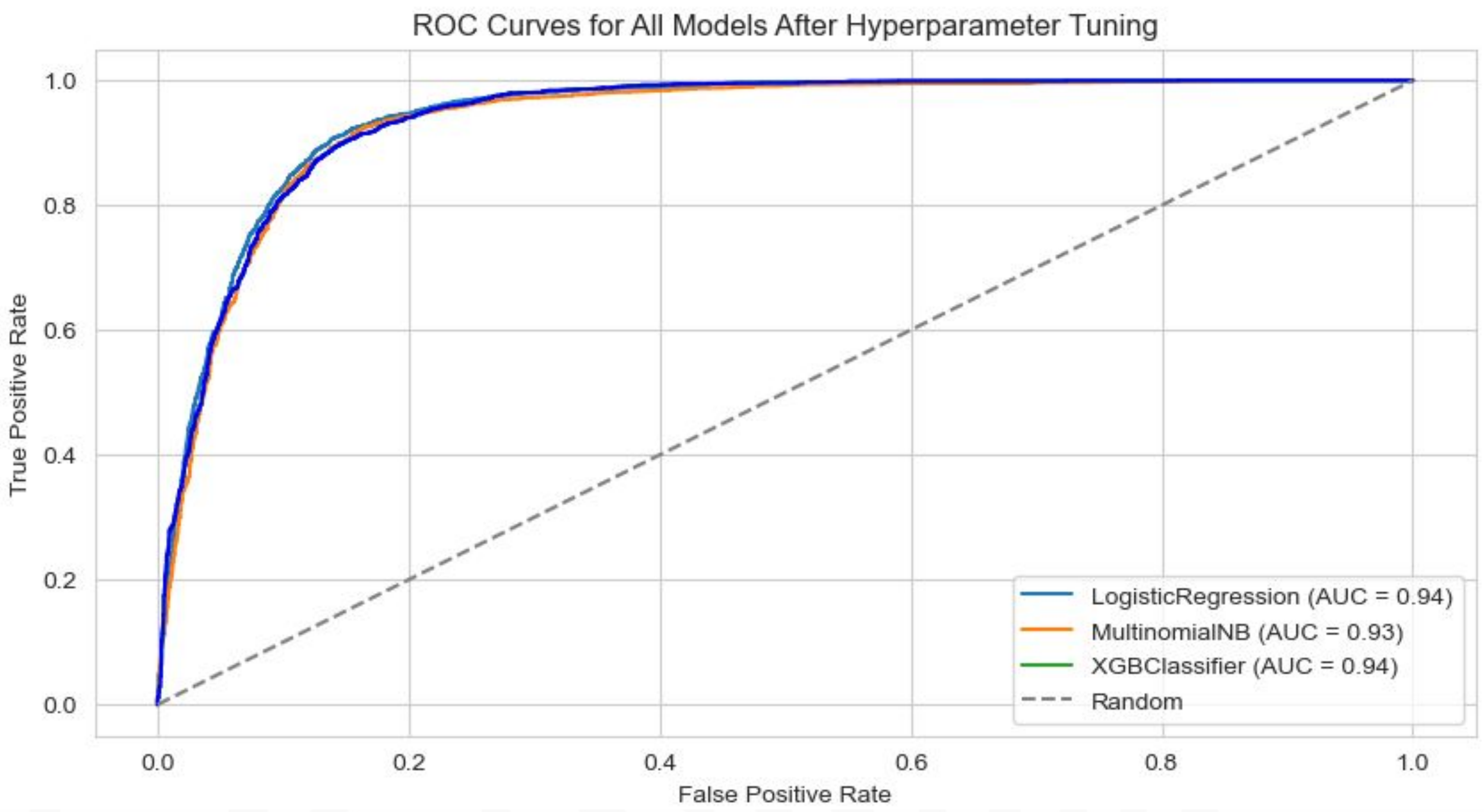
Find the best parameters for the vectorizer and classifier simultaneously with optimization using F1 Score scoring.

	Parameter	Range
LogisticRegression		
Regularization Strength	C	[0.001, 0.01, 0.1, 1, 10, 100]
Norm of Penalty	penalty	[None, 'l1', 'l2']
Optimization Algorithm	solver	['liblinear', 'newton-cg', 'lbfgs', 'sag', 'saga']
MultinomialNB		
Smoothing Frequency	alpha	range(-10,10)
Prior Probability	fit_prior	[True, False]
XGBClassifier		
gradient descent adjustment	learning_rate	[0.05, 0.01, 0.1, 0.5]
Number of iterations	n_estimators	[10,50, 100, 200, 300,]
Minimum instance weight in a child	min_child_weight	[1e-3, 1e-2, 1e-1, 1, 10]

Modeling and Evaluation (Hyperparameter Tuning)

Evaluation Test

Classifier	F1 Score	ROC AUC
LogisticRegression	0.868952	0.941922
MultinomialNB	0.864235	0.934770
XGBClassifier	0.851917	0.939306



Logistic Regression model has better F1 Score and ROC AUC

Modeling and Evaluation

Classifier	F1 Score Base	F1 Score HP	F1 Score Difference
LogisticRegression	0.862964	0.868952	0.005989
MultinomialNB	0.864439	0.864235	-0.000204
XGBClassifier	0.857849	0.851917	-0.005932

Classifier	Precision_Score_base	Precision_Score_HP	Precision_Difference
LogisticRegression	0.851669	0.850168	-0.001501
MultinomialNB	0.834360	0.835333	0.000972
XGBClassifier	0.837477	0.851089	0.013612

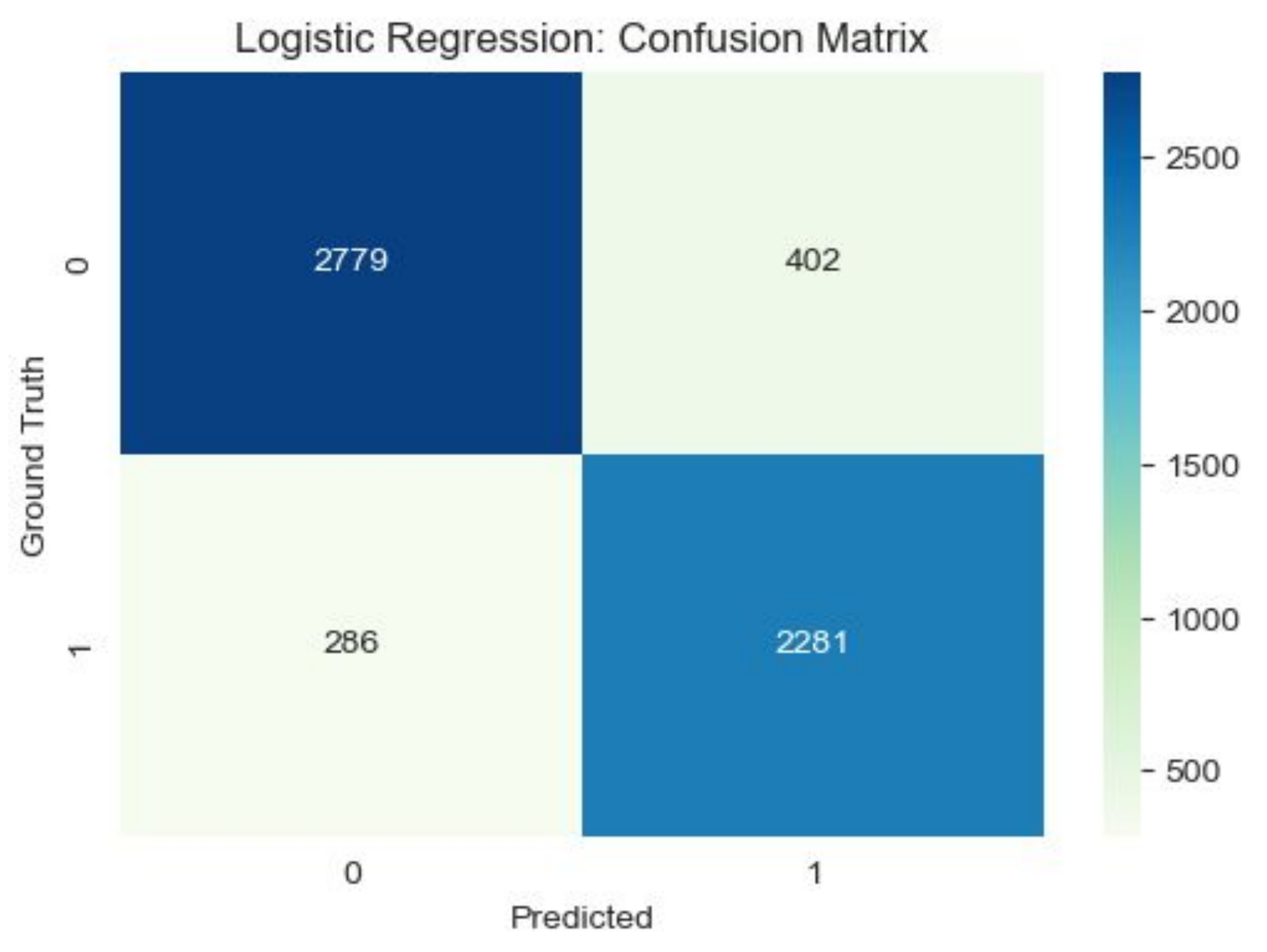
Classifier	Recall_Score_base	Recall_Score_HP	Recall_Difference
LogisticRegression	0.874562	0.888586	0.014024
MultinomialNB	0.896767	0.895208	-0.001558
XGBClassifier	0.879236	0.852746	-0.026490

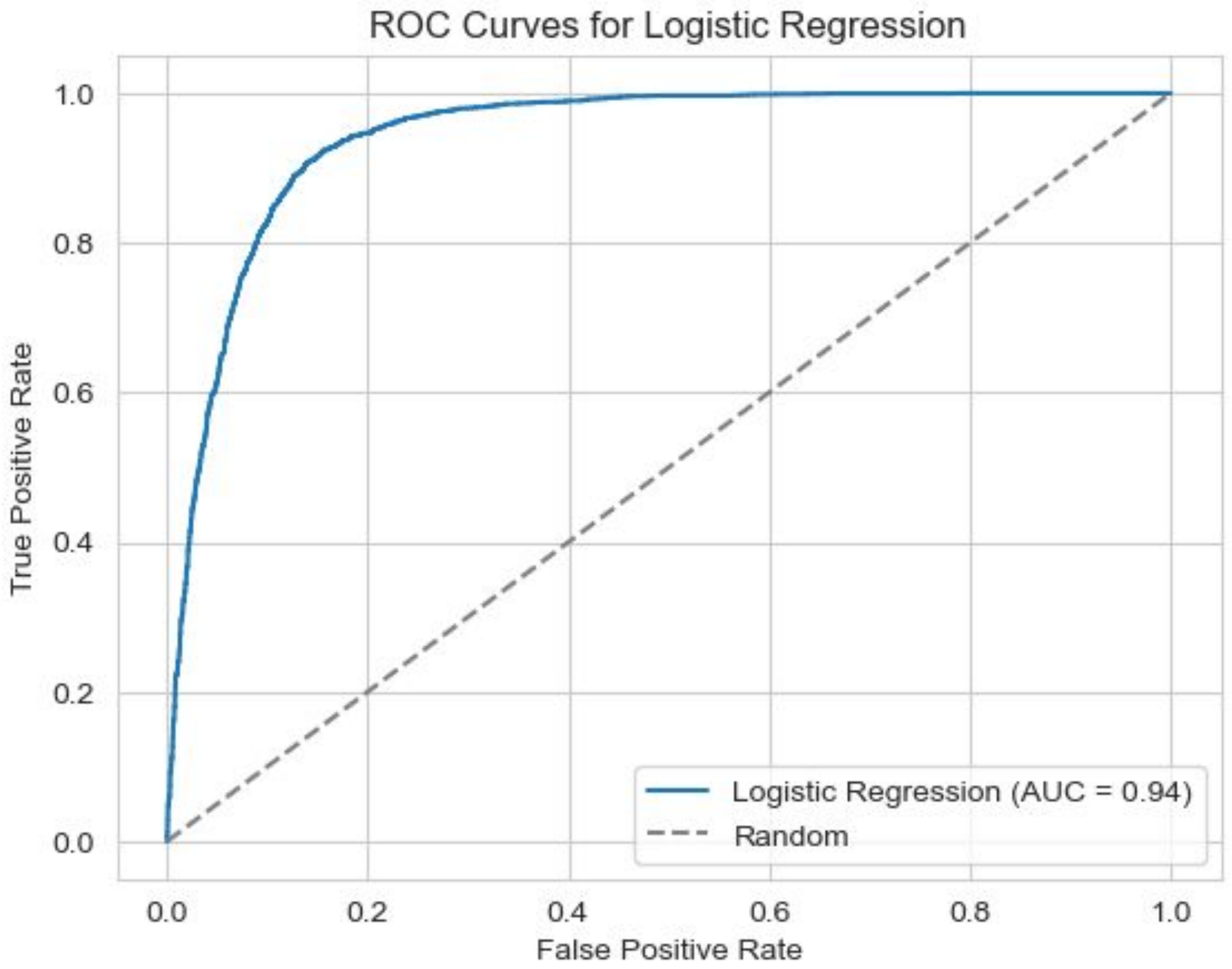
Comparation Base Model and Hyperparameter Tuning

Logistic Regression

Has a significant increase in Recall so that the F1 Score is higher than other models, which means that the logistic regression model is able to predict negative sentiment better.

Modeling and Evaluation (Final Model)





Best Model: Logistic Regression (F1 Score : 0.868952)

Best Parameter LogisticRegression :
'C': 1, 'model__penalty': 'l2', 'solver': 'saga'

Best Parameter TFIDF Vectorizer :
'max_features': None, 'ngram_range': (1, 2)}

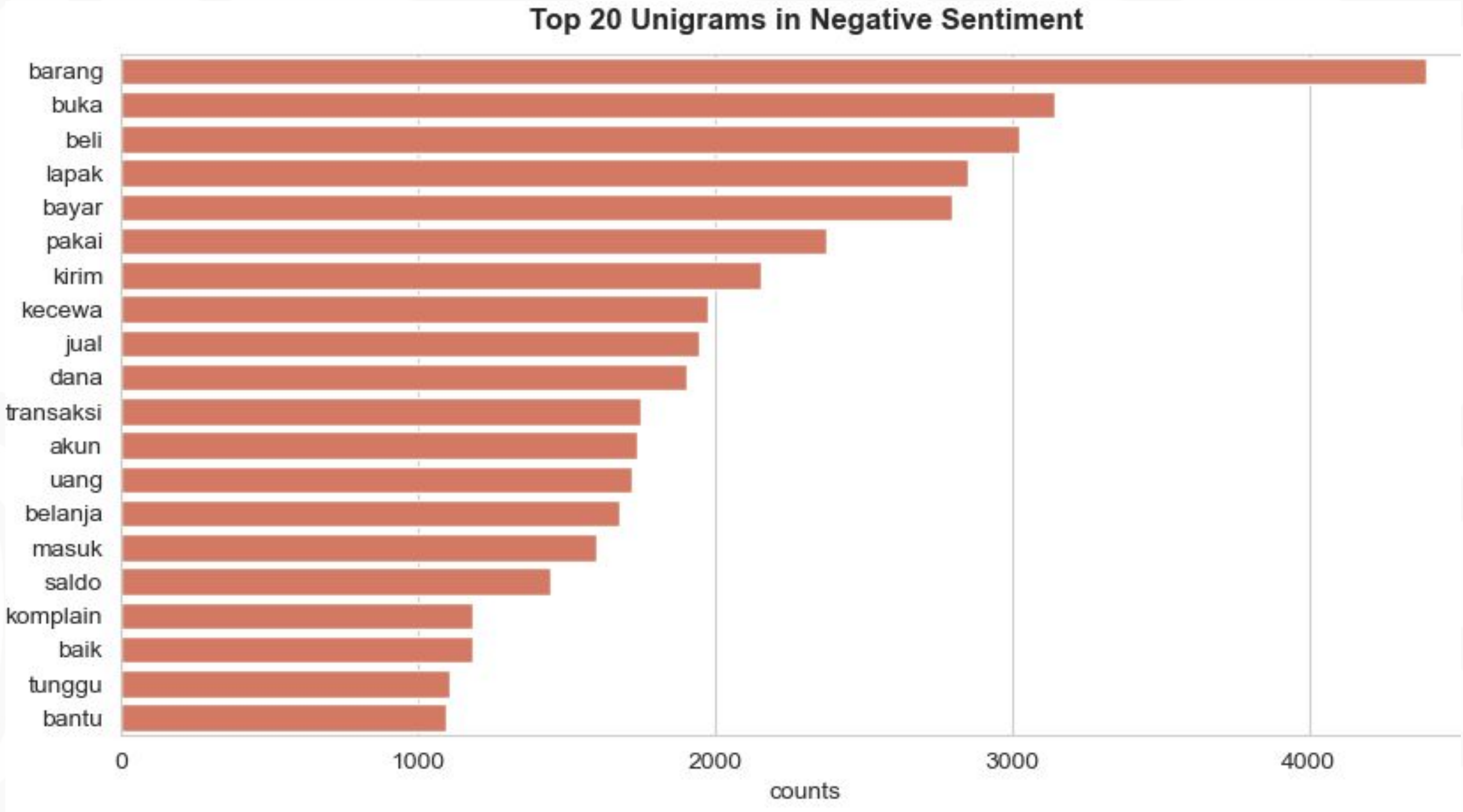
Deep Dive Negative Sentiment Analysis

Deep Dive Negative Sentiment Analysis

N-grams using CountVectorizer

Used to classify sentiment by creating clusters of text.

Divided into 3 groups (Unigrams, Bigrams, Trigrams)



Words that often appear in negative reviews (gram = 1)

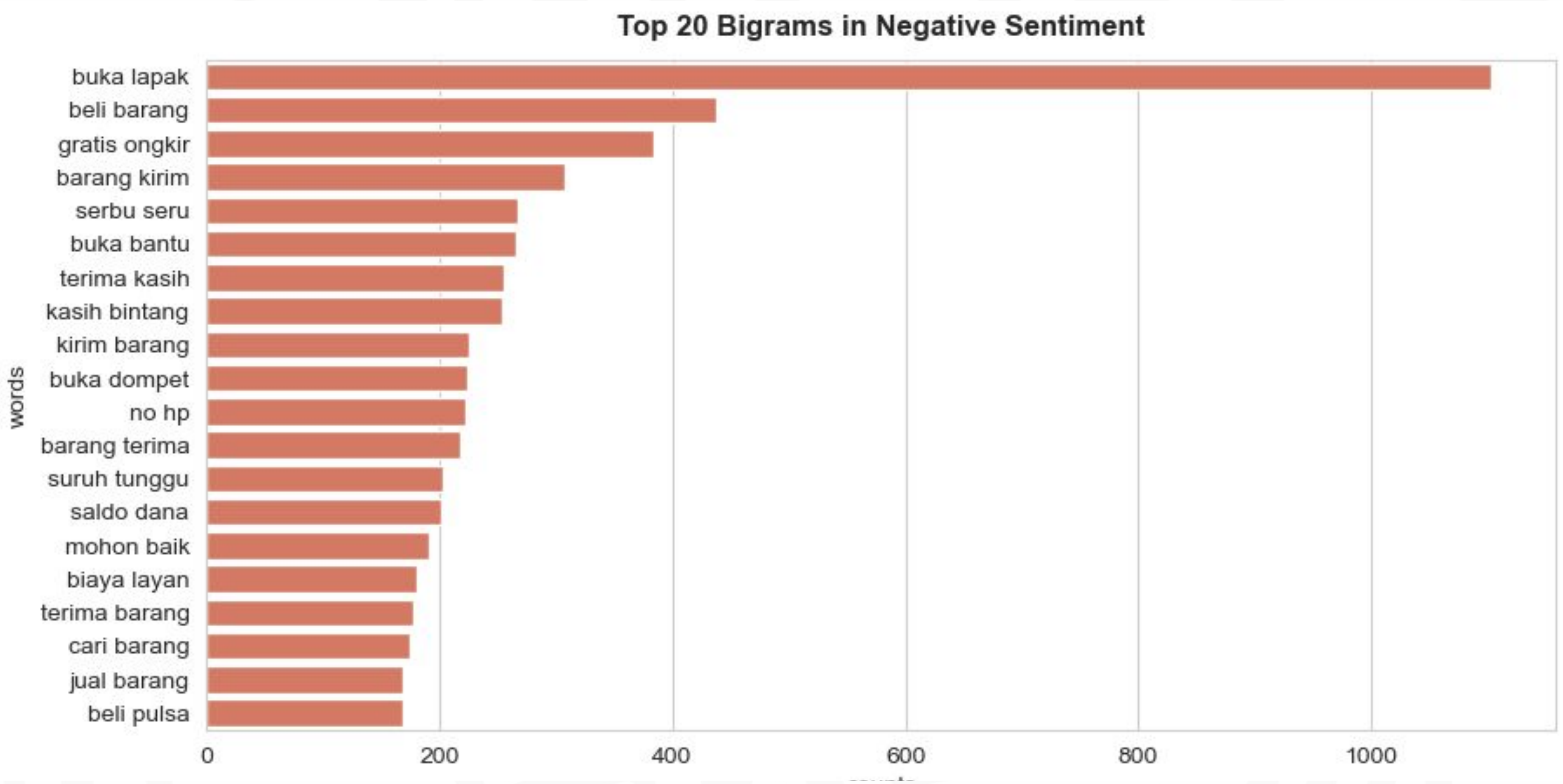
- The word **"barang"** that frequently appears in the negative category in sentiment analysis

Sample User Reviews:

"si paling ribet ribet buat para seller mau update ganti gambar aja susah tidak bisa diganti jika **barang** sudah pernah laku dengan alasan katagori tidak sesuai padahal yang merubah katagori nya pihak aplikasi nya sendiri satu satunya jalan ya di hapus bikin lagi yang baru terimakasih"

- The word **"buka"** is the most frequently occurring word in the negative category after the word **"barang"**.
- The word **"beli"** is the most frequently occurring word in the negative category after the words **"barang** and **buka"**.

Deep Dive Negative Sentiment Analysis



Conclusion:

- App users experience difficulties in paying through digital wallets
- Damaged goods are returned to the customer but the shipping costs are not reimbursed (this can be considered for platform owners)
- Lack of engagement with customers such as providing free shipping or discounted shipping vouchers when other online shops or e-commerce are massively providing vouchers to increase loyalty.

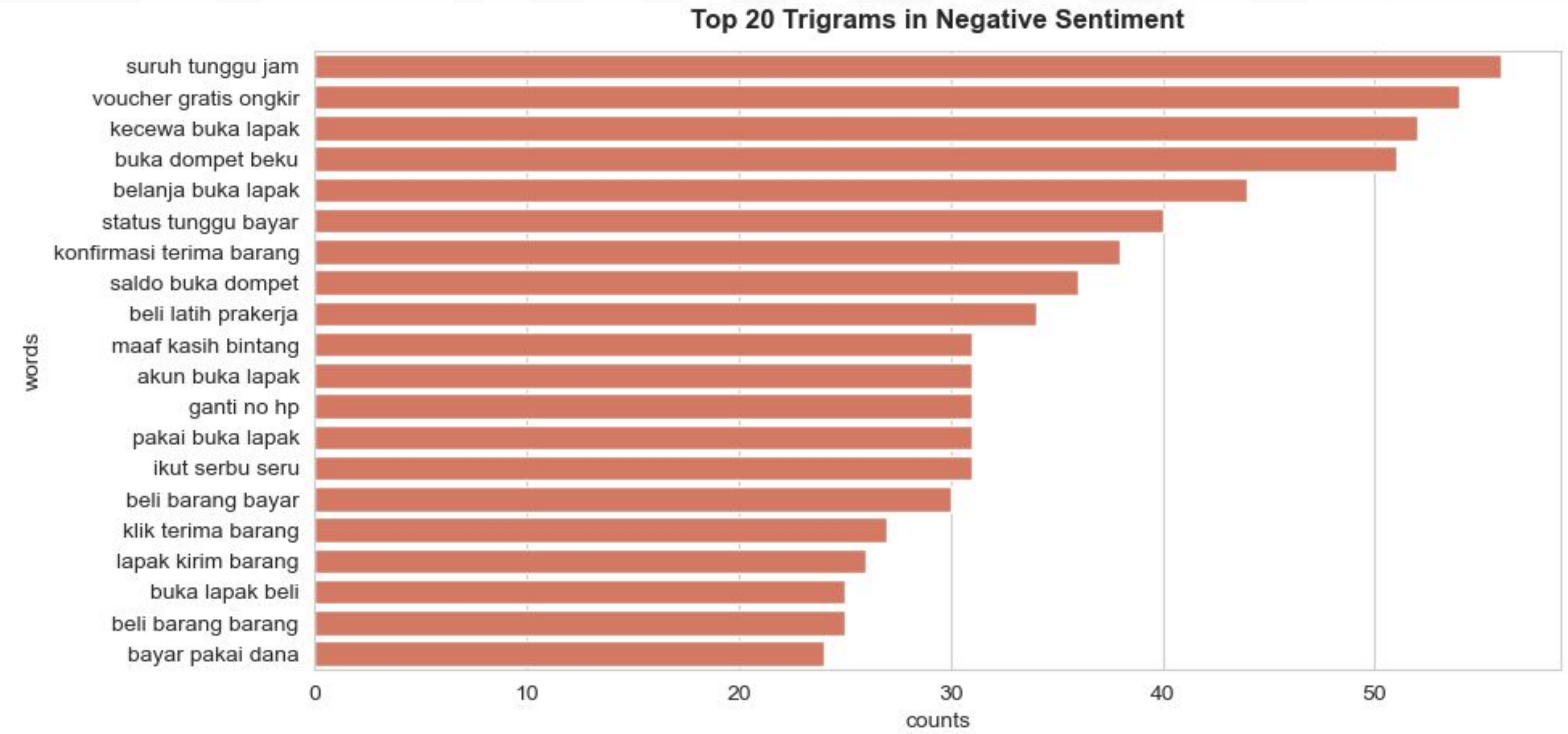
Words that often appear in negative reviews (gram = 2)

”buka lapak, beli barang, gratis ongkir”

Sample User Reviews:

- “**buka lapak** dompet digital dana ribet bayar”
- “baik atur komplain rugi **beli barang** rusak tidak fungsi uang dibalikin beli kembali bareng rugi ongkir kesini berkat konsumen beli pulsa biaya layan harga pulsa untung”
- olshop miskin pelit **voucer gratis** ongkir beli pakai ongkirnya potong beli item game pajak karuan pasar masuk akal bangkrut“

Deep Dive Negative Sentiment Analysis



Conclusion:

- App users complain about the waiting time service (this can be an important concern for platform owners)
- App users complain about the lack of free shipping vouchers, even though this can be one of the strategies to increase engagement with users so that they are loyal to use the application so that it does not lose competitiveness with other platforms.

Words that often appear in negative reviews (gram = 3)

“”suruh tunggu jam, voucher gratis ongkir”

Sample User Reviews:

- “sudah transfer via bca virtual acount transaksi kadaluarsa padahal sudah transfer saldo di rekening bca sudah terpotong beberapa kali komplain **suruh nunggu** terus sekarang sudah lebih dari **jam** tiap komplain jawaban seperti itu saja lebih baik belanja di tokopedia lebih amanah”
- “dulu nih aplikasi bisa bersaing dengan aplikasi belanja online lainnya ehh malah sekarang pelit **voucher gratis ongkir** padahal dulu selalu ada **voucher gratis ongkir** ”,
- 'good by bukalapak sedih makin lama makin pelit **voucher gratis ongkir** mari kita pindah ke lazada shopee'

Main Business Problem

Based on N-grams Analysis

- The current application is still not user friendly because there are still customers who experience difficulties and there is no information provided if there are items or products that are removed from the platform for sellers.
- The service provided is still not optimal in handling return or refund problems
- Lack of engagement with customers because the majority of negative sentiments given on average complain about the absence of vouchers or free shipping promos. Even though this can increase engagement with customers

Business Recommendation

Periodically evaluate the services and information provided to customers and improve the quality of applications to make them more user friendly.

Provide vouchers or exciting promos according to customer segments to increase engagement and loyalty in using the application.

Conclusion

Conclusion

<u>Dataset</u>

positive sentiment is higher than negative sentiment but the difference is only about +/- 22%.

<u>Best Model</u>

Logistic Regression and TFIDF with Hyperparameter Tuning get **F1 Score of 0.868952 or 86.89%**

Positive sentiment has the most words

"terima kasih, mudah, paham, bantu, latih, prakerja"

Negative sentiment has the most words

"kecewa, bayar, pakai, komplain, barang, buka, beli"

<u>Business Recommendation</u>

To improve the quality and service in providing information and also provide vouchers or exciting promos according to the customer segment to increase engagement and loyalty of application users.

THANK YOU !

Lets Connect

Gmail Github Linkedin