ProDev: Group 3 Project Report

M-Pesa Sentiment Analysis (M-Pesa) Using Twitter Data

Group Members
Eva Gitau
Eric Cheruiyot
Colleta Nandutu
Samwel Omondi
Miriam Wangari
Lisa Kangendo

February 2023

BUSINESS UNDERSTANDING

1.0 Project Overview

Sentiment analysis is a technique used to determine and categorize an opinion about a product, service or idea as positive, negative or neutral. Sentiment analysis helps organizations gather insights from unorganized and unstructured text that comes from online sources such as emails, blog posts, or social media channels. It is often performed to help businesses monitor brand and product sentiment in customer feedback and understand customer needs. Sentiment analysis is leveraged to identify the polarity of information (positive vs. negative), emotion (anger, happiness, sadness, etc.), and intention (e.g., interested and not interested).

This project aims to analyze and accurately predict a tweet's sentiment about M-pesa mobile money services in Kenya so as to understand the general public opinion and make improvements on M-pesa products and quality of service.

M-Pesa is a mobile money service that was first launched in Kenya in 2007. It allows users to store and transfer money, pay bills and purchase goods and services using their mobile phones. M-Pesa is currently the most popular alternative to traditional banking in many African countries, especially Kenya, where a large proportion of the population is unbanked. The service has expanded to other countries in Africa, Asia, and Europe, and has become a model for other mobile money services around the world. M-Pesa has been credited with revolutionizing mobile banking and transforming the way people in developing countries access financial services.

1.2 Problem Statement

The widespread adoption of M-pesa services among the online users has elicited mixed reactions on the products and quality of service. Our project's aim is to perform sentiment analysis on tweets related to M-pesa mobile money service and predict a tweet's sentiment accurately in order to understand the customers' opinion on the brand, improve on quality of the products, delivery of service and brand expansion.

1.3 Research Objectives

1.3.1 Main Objective

To understand customers' satisfaction with Mpesa products and services and gauge whether a tweet's sentiment is neutral, positive or negative in order to improve the quality of service

1.3.2 Specific Objectives

- To gauge public opinion on M-pesa so as to assess m-pesa brand reputation.
- To identify potential causes for positive or negative sentiment towards M-pesa.
- To identify the most common topics and hashtags related to M-pesa.
- To identify the most common sentiment related to m-pesa.

- To predict a tweet's sentiment on m-pesa accurately.

1.4 Specifying the Question

What are the sentiments expressed in tweets related to m-pesa and can we predict a tweet's sentiment accurately?

1.5 Metrics of Success

The sentiment analysis will be considered successful when we are able to accurately classify and predict the sentiment expressed in tweets related to mobile money service m-pesa, with a focus on high accuracy, precision and recall.

1.6 Experimental Design

1. Defining the research question.
2. Scraping, loading the data and previewing preliminary characteristics of the dataset.
3. Data cleaning and preprocessing with EDA.
4. Converting the tokenized tweets into numerical data.
5. Splitting the data into training and testing sets.
6. Fitting a machine learning model to the training data.
7. Evaluating the model's performance on the testing data.
8. Deployment.
9. Conclusion.
10. Recommendations.

1.7 Relevance of the data

The sentiment expressed in tweets related to M-pesa mobile money service is relevant in understanding the public opinion towards the service. This information can be useful for the service provider in improving their service, addressing any negative sentiment expressed by users and to research real customer needs and assess brand reputation.The tweets data was scraped from twitter between the dates of "2018-01-01" and "2023-02-09".

2.0 Data Understanding

Here we the different aspects of the scrapped dataset, such include
- the size of the dataset,
- the types of variables in the dataset,
- the descriptive statistics of the variables and
- the detection of errors, outliers and missing values

2.1 Data Collection

The data was extracted from twitter website that contain the preset hashtags related to M-Pesa.

2.2 Data Scraping

This code is a script to scrape tweets containing hashtags related to mobile money service m-pesa. The hashtags to be searched are defined in the tags list. The time frame for the tweets to be scraped is set between since_time and until_time.

The script makes use of the multiprocessing library to run multiple scraping processes in parallel. The function scrape(tag) is called for each hashtag in the tags list. The function scrapes tweets containing the hashtag using the TwitterSearchScraper class from the snscrape.modules.twitter library. The resulting tweets are then appended to a list tweets_list2 and saved to a data frame df1. The data frame is then saved as a csv file in the "twitter_data" folder.

Finally, the script creates a data frame data by reading in all the csv files in the "twitter_data" folder using pd.read_csv and concatenating the resulting data frames using pd.concat.

2.3 Data Description

The dataset extracted from twitter has 21768 rows and 4 columns. The rows represent tweets and each tweet has 4 features.

2.4 Data Quality Verification

Checked for missing values in our dataset which we went ahead and dropped.

The dataset has 4 columns, namely:

- Tweet
- UserName
- Id
- Unnamed:0

We dropped the tweetid, UserName and Unnamed:0 columns since they are not needed for our analysis.

We went ahead and checked for duplicates. The dataset had 149 duplicates which we dropped.

We use the TextBlob library in Python, which uses a pre-trained sentiment analysis model to define a function called "sentiment" that takes tweets and returns the sentiment polarity of the tweets using the TextBlob library. The sentiment polarity is a value between -1 and 1 that represents the sentiment of the tweet, with -1 being negative, 0 being neutral, and 1 being positive.

You can then use this function to create a sentiment column in a Pandas DataFrame by applying the function to each element in a column containing the tweet data.

Created a new column called sentiment that contains the sentiment polarity of each tweet and another column called sentiment_label that contains a label for each tweet as positive, negative, or neutral.

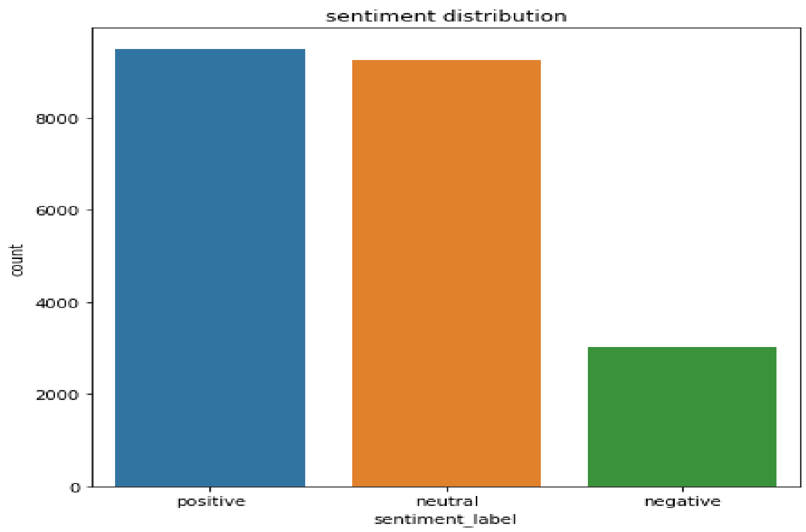
	Tweet	sentiment	sentiment_label
0	The best that happened in Kenya is #Mpesa. \nE...	0.750000	positive
1	The CS\n#Safaricom\n#Mpesa\nNjugush\nYaya\nPre...	0.000000	neutral
2	Here we enjoy the Sweet & Sour nature of I...	0.346875	positive
3	But since its from Mpesa and I sent the revers...	-0.200000	negative
4	Congratulations to the Central Bank of Kenya o...	0.150000	positive
...	...	...	...
21761	@SafaricomLtd Review your rates...failure to t...	-0.062500	negative
21762	Well Come To safaricom. \nWhere checking your...	0.450000	positive
21763	@KCBGroup @KCBCare Good afternoon, I wanted to...	0.700000	positive
21764	slay queens ends their posts by saying~\nfollo...	0.000000	neutral
21765	Sooooooo many new year messages. Amen . The re...	0.335227	positive

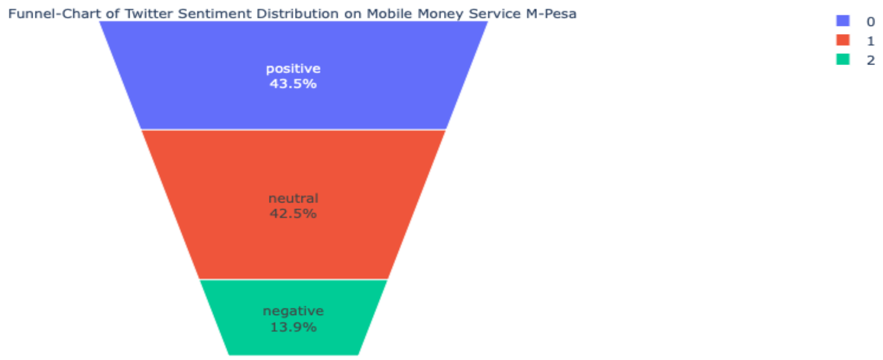
21764 rows × 3 columns

3.0 Data Exploration

3.1 Sentiments distribution

There are 9477 positive m-pesa tweets, 9257 neutral and 3,035 negative m-pesa sentiment tweets.

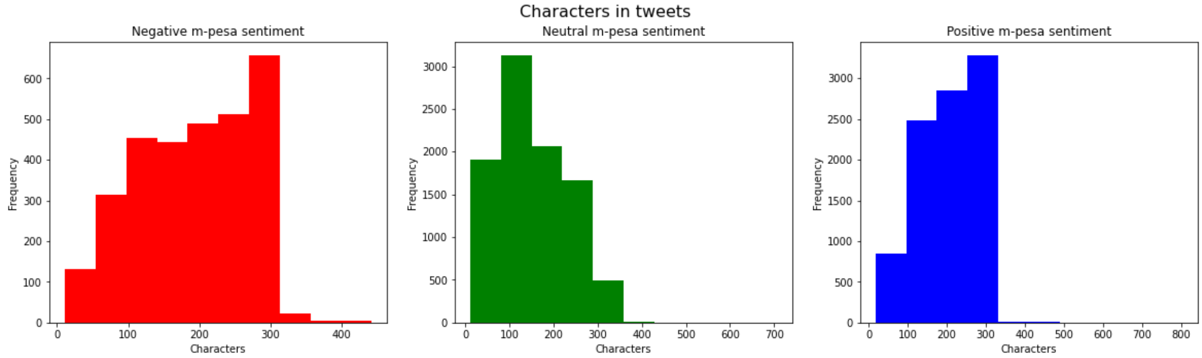




Most tweets are positive sentiments on mobile money m-pesa services at 43.6% . This is closely followed by category 1, tweets which are neutral at 42.5% and taking the last category, 2, are tweets which have negative sentiments on m-pesa services at 13.9%.

3.2 Investigating the number of characters, words in the tweets.

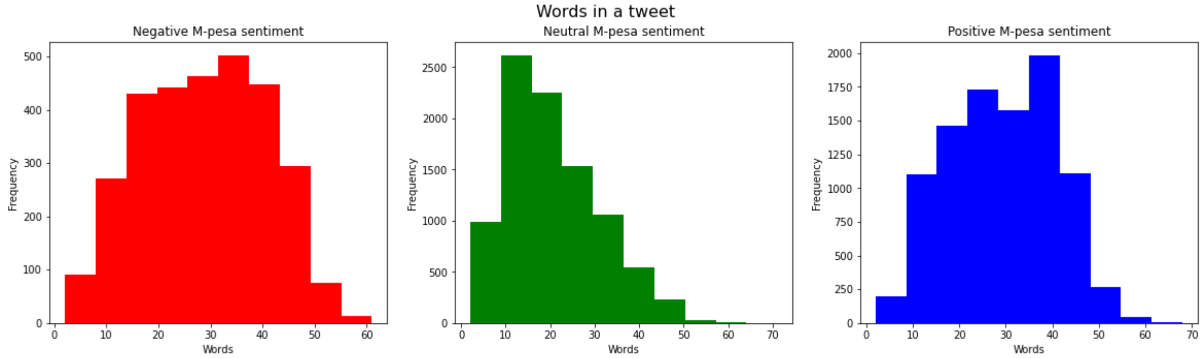
We plotted histograms of distributions of the number of characters in tweets.



From the visualizations above, the characters in a tweet range from 0 to about 420 for negative sentiment, while those of neutral tweets range between 0 to about 450 characters, positive sentiment tweets have about 0 to about 500 characters. In general, most tweets range between 100 and 300 characters across all sentiments.

3.3 Investigating number of words in tweets

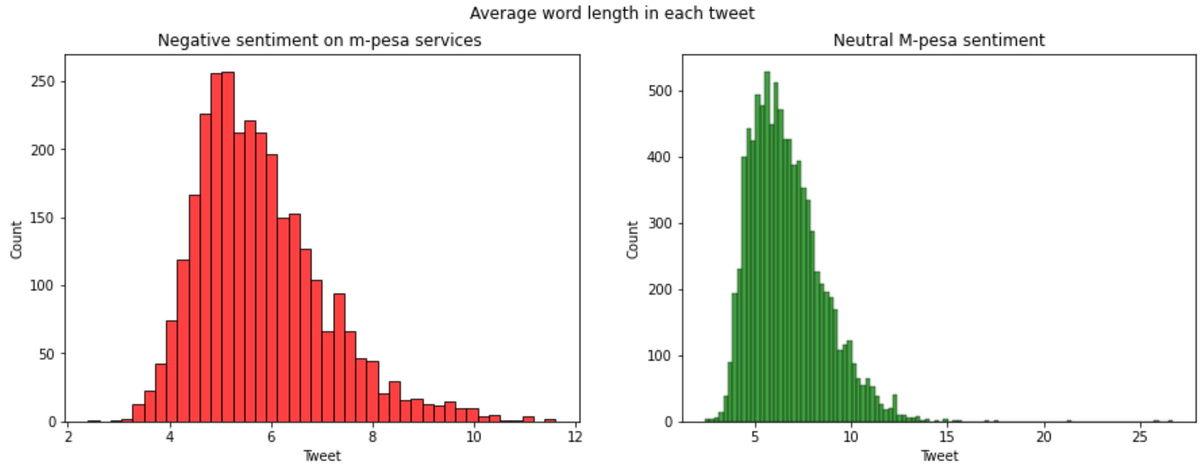
We plotted histograms of distributions of number of words in tweets.

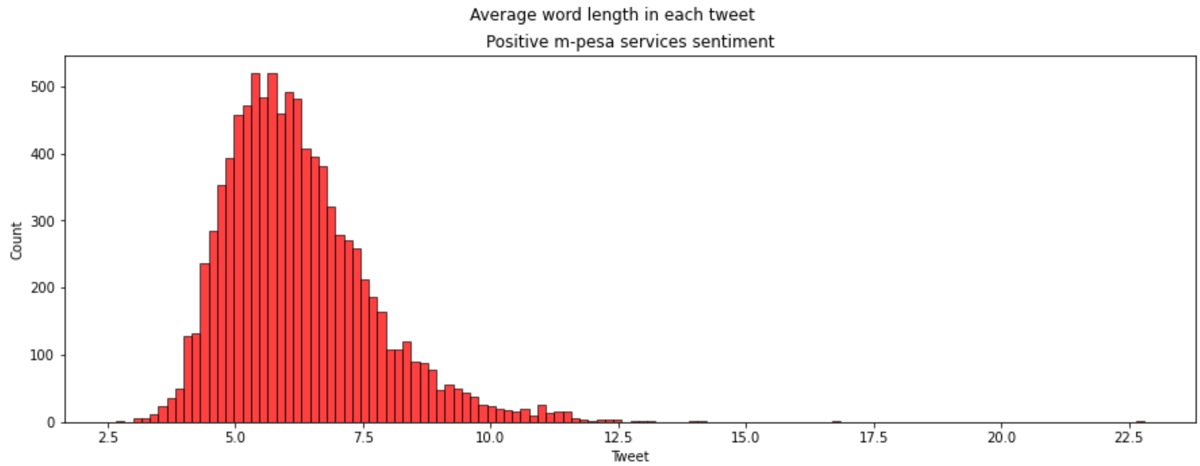


From the visualizations above, the words in a tweet range from 0 to about 60 for negative sentiment, while those of neutral tweets range between 0 to about 60 words, positive sentiment tweets have about 0 to about 70 words. In general, most tweets range between 10 and 50 words across all sentiments.

3.4 Average word length

We plotted histograms of distribution of average word length in tweets.





From the visualizations above, the words in a tweet range from 5 to about 10 characters long on average in general across all sentiments.

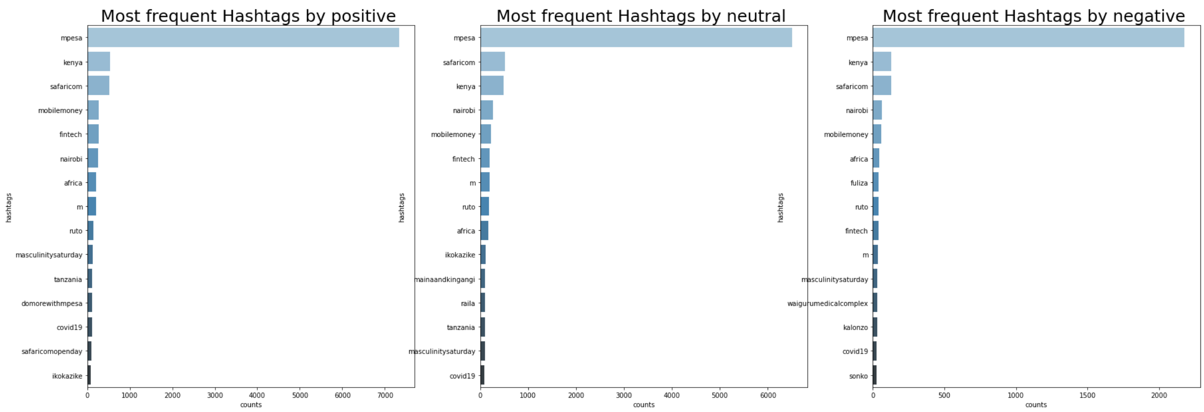
3.5 Tweet hashtag summary

Since Twitter uses Hashtags almost like a summarization feature (at least in the sense of highlighting core ideas). We look at some of top hashtags for each of the classes of sentiment. We'll then make "word clouds" to visualize their prominence.

Table of top 15 hashtags

	negative		neutral		positive	
	hashtags	counts	hashtags	counts	hashtags	counts
0	mpesa	2184	mpesa	6498	mpesa	7339
1	kenya	130	safaricom	515	kenya	527
2	safaricom	126	kenya	485	safaricom	517
3	nairobi	64	nairobi	269	mobilemoney	275
4	mobilemoney	60	mobilemoney	232	fintech	265
5	africa	42	m	197	nairobi	252
6	fuliza	41	fintech	195	africa	206
7	ruto	41	ruto	182	m	200
8	fintech	38	africa	174	ruto	141
9	m	34	ikokazike	121	masculinitysaturday	126
10	masculinitysaturday	30	mainaandkingangi	106	tanzania	119
11	waigurumedicalcomplex	29	raila	103	domorewithmpesa	105
12	kalonzo	29	tanzania	102	covid19	105
13	covid19	26	masculinitysaturday	96	safaricomopenday	104
14	sonko	24	covid19	93	ikokazike	86

Visualizing the frequency of hashtags in barplots for each sentiment categorization



The most popular hashtags are, broadly, mpesa, safaricom and Kenya. Which is expected, given the topic; but also, among the top 10 is a hashtag relating to Ruto probably his 2022 presidential campaign .

The fuliza hashtag refers to a 2019 overdraft service that m-pesa launched on it's mobile money wallet in conjunction with NCBA bank. Here it features quite frequently on the negative sentiment tweets.

Africa featuring in all the sentiment categories shows that m-pesa is indeed popular in the continent. Tanzania particularly in positive and neutral sentiment categories shows that m-pesa enjoys popular appeal and after kenya that might be the next largest market for m-pesa.

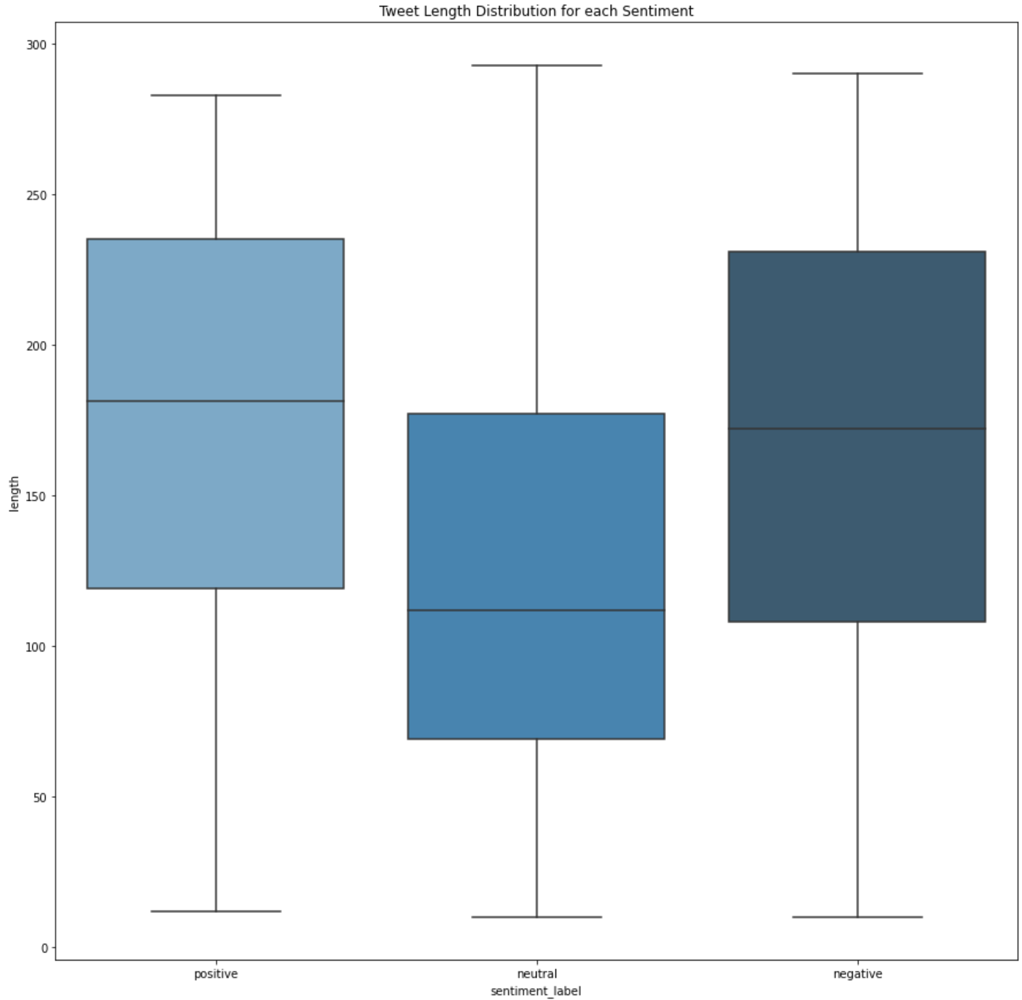
Interestingly, covid19 made the shortlist of the positive sentiment. This is likely attributed to m-pesa to bank transaction charges/fees being removed during the pandemic period.

4. 0 Data Processing

4.1 Final Data Cleaning and Pre-processing

As part of preparing the data for modelling, we defined a function that removes duplicates from the tweet messages, performed lemmatization and removed stopwords from the tweets.

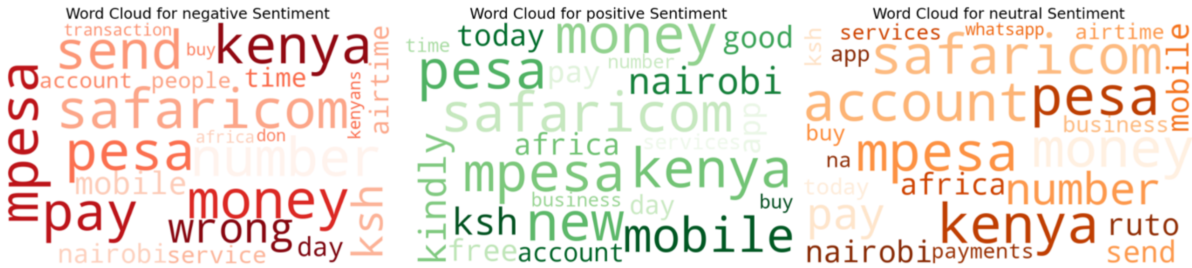
Visualizing the average length of tweets after some transformation as above
We plotted boxplots of average tweet lengths.



From the visualization above, it is evident that most tweets range from 50 to 250 characters while some are in excess of 250 characters.Other tweets are observed to be very short, essentially less than 50 characters

4.2 Identifying the frequent words/Buzzwords

Building a wordcloud visualizer for each word frequency in the sentiments.



The top 4 buzzwords are mpesa, pesa, kenya and safaricom . This seems to indicate that a lot of the same information is being shared/viewed – this applies across all sentiments. While we can't conclude that's a result of the "filter bubble", it certainly seems like that might be a latent (hidden) cause.

Interestingly, 'ruto' occurs in the neutral sentiment tweets. This may not be surprising given his political activity during the timeframe the tweets were scraped – this is something that likely warrants further investigation especially along this axis of Neutral.

The word 'wrong' occurs quite frequently in the negative sentiment. This might be attributed to tweets where users tweeted about wrong number transactions which are obviously annoying hence the negativity of the sentiment.

Take a look at the table above, you'll see that 'free' actually shows up in the positive sentiment quite frequently. This would imply that customers tweeted quite frequently about their satisfaction with the free transactions during the pandemic

4.3 Final Dataframe

5.0 Modeling

5.1 Implementing the Solution using Long Short Term Memory (LSTM) in Keras

The dataset is split into training and test sets, tokenized using the Tokenizer function from Keras, and sequences are padded to the same length. The LSTM model is then defined with an embedding layer, spatial dropout, LSTM layer, and output layer with softmax activation. The model is trained for 10 epochs with a batch size of 30, and its performance is evaluated using accuracy, precision, recall, classification report, and confusion matrix. Finally, the loss and accuracy of the model are plotted using matplotlib.

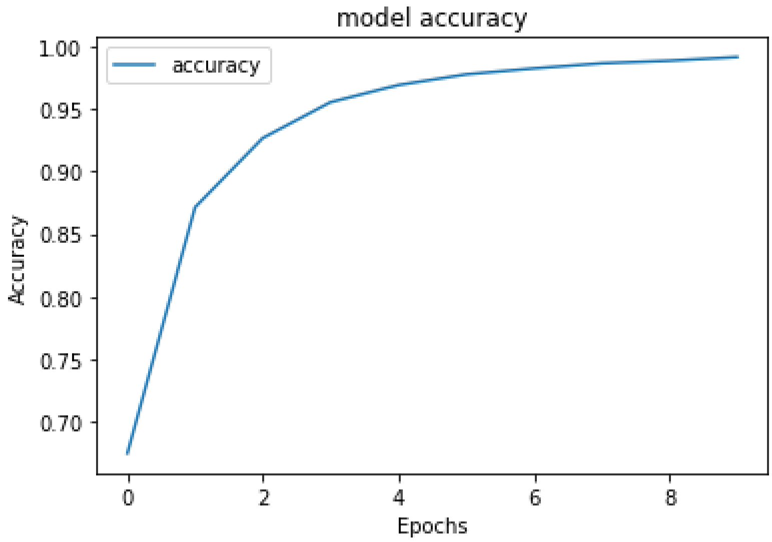
```
              precision    recall  f1-score   support

           0       0.91      0.87      0.89      1915
           1       0.72      0.71      0.71       575
           2       0.85      0.90      0.88      1864

    accuracy                           0.86      4354
   macro avg       0.83      0.82      0.82      4354
weighted avg       0.86      0.86      0.86      4354

---------------------------------------------------------------------
[[1663   71  181]
 [  66  406  103]
 [ 104   89 1671]]
```

5.2 Accuracy of the model



LSTM has an accuracy score of 86% at 10 epochs.

5.3 Challenging the Solution

We challenged the solution using XGBoost.

	precision	recall	f1-score	support
0	0.65	0.70	0.67	1915
1	0.49	0.07	0.13	575
2	0.60	0.71	0.65	1864
accuracy			0.62	4354
macro avg	0.58	0.49	0.48	4354
weighted avg	0.61	0.62	0.59	4354

Accuracy of our model is 0.6196600826825908

XGBoost gave us an accuracy of 62%. A look at the scores indicates that LSTM model performs better, thus we shall proceed with LSTM to deployment.

6.0 Deployment

The deployment is done using streamlit. We saved the trained LSTM model to an h5 file in our working directory. We did pre-processing of the input text before deployment testing using the saved tokenizer json file and lemmatized removing stopwords and unwanted characters before fitting the model and making our predictions .Deployment github

7.0 Conclusions.

- Most tweets are positive sentiments on mobile money m-pesa services at 43.4%, closely followed by tweets which are neutral at 42.6% and the last are tweets which have negative sentiments on m-pesa services at 14%.
- COVID-19 had a positive impact on M-Pesa transactions due to the removal of transaction fees during the pandemic.
- M-Pesa is a popular mobile money wallet in Kenya and Tanzania, and it has a positive sentiment among users in those countries.
- The most popular hashtags related to M-Pesa are mpesa, safaricom and Kenya.
- Users tweet frequently about the Fuliza overdraft service, and it features quite negatively though popular.
- Users tweet frequently about the transactions that ended up in money being sent to wrong numbers and these tweets are mainly negative sentiment-wise.

8.0 Recommendations.

- M-Pesa could leverage the positive impact of COVID-19 on its public image by the removal or reduction of transaction fees which attracted positive public sentiment during the pandemic.
- Since the Fuliza overdraft service is frequently mentioned in negative sentiment tweets, M-Pesa and NCBA bank could work on improving the service to increase customer satisfaction.
- Further investigation could be conducted on the relationship between Ruto and neutral sentiment tweets to gain a deeper understanding of the role of politics in customer sentiment towards M-Pesa though we suspect it was spam.
- M-Pesa could analyze tweets related to "wrong" number transactions and work on improving the user experience to reduce negative sentiment. This can be done by ensuring the users can verify the correct number to transact with and facilitating easier reversal of erroneous transactions for better user satisfaction.
- Safaricom could leverage the positive and neutral sentiment in Tanzania to market M-pesa more and expand its reach in the market which is larger than Kenya demographically and therefore has more potential.

Reference

Collab Notebook:
https://colab.research.google.com/drive/1f4kBYGsWXY6UpnKgSDx4Dq99SPpyLiiT?usp=sharing

GitHub Repository: Deployment github