**Azubi Hackathon 2020**

**RAM TEAM (Group 6)**

**R**azack Abubakar

**A**ma Owusu- Darko

**M**ichael Kwame Sekyere

**Abstract**

In recent years, the interest among the research community in sentiment analysis (SA) has

grown exponentially. It is only necessary to see the number of scientific publications and

forums or related conferences to understand that this is a field with great prospects for

the future. On the other hand, the Twitter boom has boosted investigation in this area

due fundamentally to its potential applications in areas such as business or government

intelligence, recommender systems, graphical interfaces and virtual assistance. However, to

fully understand this issue, a profound revision of the state of the art is first necessary. It

is for this reason that this paper aims to represent a starting point for those investigations

concerned with the latest references to Twitter in SA, in the political atmosphere of Ghana.

***Assumptions/Hypotheses***

1. Social media cannot fully replace the traditional media in politics
2. Followers of political leaders on twitter is a good estimate for the number of party sympathizers
3. Social media handles are operated directly by political leaders in Ghana
4. Positive tweeting political leaders are hugely admired and win elections

***1.0 Introduction***

Ghana, the ever-shinning black star of Africa. Ghana, the beacon of hope of democracy in the west African region. Ghana, the first sub Saharan country to gain her independence from the western world.

Accolades, attributes, achievements, milestones and pacesetting records of this west African country with a population just over 30 million people cannot be overestimated. The struggles, the toil, the humiliation, the torture and murder of her founding fathers and brave women paid off finally on the 6th day of March, 1957 when Dr. Kwame stood on the podium at the independence square, now called the black star square to declare this young and promising nation, free from colonial rule, free from having to answer western powers, free from shackles of colonization and free forever.

The history of this proud nation has not always been smooth or glorious. She has had its fair share of government overthrown and military interventions. Since her first democratic elections in 1960, which saw Dr. Kwame Nkrumah declaring the country a one-party state. This first republic could only last for so long as there was growing concern about the democratic state, the people hoped to enjoy after independence

Since the overthrown of Dr. Nkrumah’s government, there have been three republics and four military interventions. The common similarity shared by this government is the approval by most of the citizenry. And anytime the masses were not satisfied by a government's performance, change becomes inevitable, be it a military intervention or at the polling station

Social media has become a powerful tool in the 21st century. Since the early 2000’s, the emergence of these social media platforms has in their own unique changed the course and ways of humanity. Social media platforms have become the most popular channels of communication thus replacing the traditional forms of media, bringing more vibrancy, low latency and the largest audience which traditional media could not account for.

Yet, until 2015 when the US president, Donald Trump launched his campaign into office, social media was not ‘really’ used to its full capacity by politicians and world leaders. To be fair to posterity, Trump’s twitter activity is what really kick start an inevitable communication platform for politicians around the world.

Trump won elections in the US and since his early days, his tweets have been considered official statements from the presidency even though they are sometimes controversial or false

With the tentacles of social media now embedded deep into politics worldwide, we used the twitter API to retrieve tweets of leading Ghanaian politicians of the major political parties in the country and performed sentiment analysis on these tweets

Sentiment analysis (also known as opinion mining or emotion AI) refers to the use of natural language processing (NLP), text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information.

The applications of sentiment analysis are broad and is widely applied to voice of the customer materials such as reviews and survey responses, online and social media, and healthcare materials for applications that range from marketing to customer service to clinical medicine. Our capstone project is about using NLP to predict the positive, negative or neutral sentiments of tweets by Ghanaian politicians and to estimate how Ghanaians interact with their political affiliates or opponents on twitter

***2.0 Data and Methodology***

***2.1 Data Collection and Pre-Processing***

More than 70,000 tweets were collected from the month of May to September in python environment using twitter search API for the two main contesting parties in the country, thus the incumbent NPP and opposition NDC

Initially a Twitter Developer account was created, but there are limitations to it. Tweets older than 7 days cannot be mined, using twitter's rest API, unless it is paid for and even when you do pay, you are restricted to a limit of 100 tweets daily. Through research we came across Optimized-modified-GetOldTweets3-OMGOT which is an improvement of the original GetOldTweets Library by Jefferson-Henrique. With this we were able to bypass the limitations of using tweepy

These tweets contained various hashtags, the popular amongst the lot being #4more4Nana and JMandJane2020. Also, the number of retweets, replies and likes on tweets by the leading figures of these political parties were also collected and recorded. These political figures have accounts named NAkuffoAddo, NPP\_GH, and MBawumia which corresponds to the incumbent president H.E Nana Addo Akuffo Addo, the official account of the National Patriotic Party, and the vice president of the country, H.E Alhaji Mahmud Bawumia respectively.

Also accounts like JDMahama, NJOAgyemang and OfficialNDCGH corresponds to the former president, H.E John Dramani Mahama and flagbearer, Prof. Naana Jane Owusu Agyemang, the running mate of Mahama and the official account of the National Democratic Congress respectively.

The data was preprocessed by dropping columns that had no bearing on our analysis and had lots of null values thus, ID, mentions, to, geo and permalinks. ID column contained an automatically ID twitter generates for a particular tweet, mentions also contained lots of null values and contained names of other tweepers refrenced or tagged in the tweet, the ‘to’ column contained other tweepers who had particular popular tweets directed to them by their friends or family and it also had lots of null values. ‘Geo’ is the location in Ghana where the tweet was sent from and it contained lots of null values

Finally, we used the collected data to do a sentiment analysis. We again preprocessed the data through these steps below

1. Clean Data: This is to remove symbols and non-alphabets that may be present in the tweet but does not give meaning to it.

2. Tokenize the data: Tokens are words separated by spaces.  Divide the texts into words or smaller sub-texts, which will enable good generalization of relationship between the texts and the labels. This determines the “vocabulary” of the dataset

3. Normalization: is done to ensure that different forms of a particular word are represented as one. For example, \*\*ran\*\*, \*\*run\*\* and \*\*running\*\* are different variation of the word. Normalization helps represent all these forms by a single form \*\*run\*\*. There are two main form of Normalization. These are

a. Stemming removes affixes from words and representing it with its root word

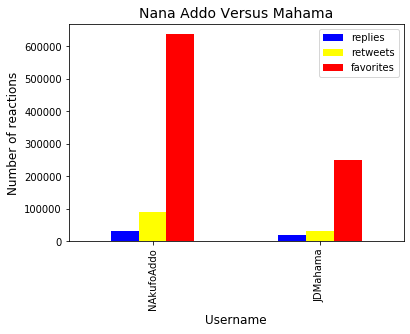
b. Lemmatization normalizes a word with the context of vocabulary and morphological analysis of words in text

4. Get sentiment: produces an output of Total Positive Tweets, Total Negative Tweets and Total Neutral Tweets

***3.0 Analysis and Insights***

***3.1 Head to Head Comparisons***

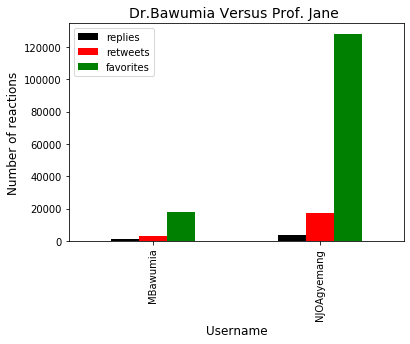
***3.1.1 H.E Nana Akuffo Addo vs H.E fmr. John Dramani Mahama***



***Graph 1***

Graph 1 is a graphical representation of replies, retweets and favorites (like reactions) to tweets by the accounts run by the leaders of the two major political parties. It can be observed that the red bar which represents the favorites column of NAkuffoAddo clearly outperforms the favorites column of JDMahama. This could be attributed to a number of reasons, the one which stands out is that NAkuffoAddo is currently the president of the nation hence has much twitter following hence reactions(retweets, replies and favorites) to his posts would be clearly much higher than the former president JDMahama. Also, NAkuffoAddo uses his official twitter account for official releases concerning the affairs of the country. Everything the former president puts out is not official and might just be his personal opinions regarding matters of the state

***4.1.2 VP Bawumia Vs Prof. Jane Naana***

***Graph 2***

Now this is an interesting graph. Graph 2 is a plot of the activity of on the account of the vice president of Ghana, MBawumia, side by side the activity against his direct competitor, NJOAgyemang. There is a huge variation in the number of replies, retweets and favorites accrued by the vice president as compared to that of the opposition, stationing herself to boot out the incumbent government out of office.

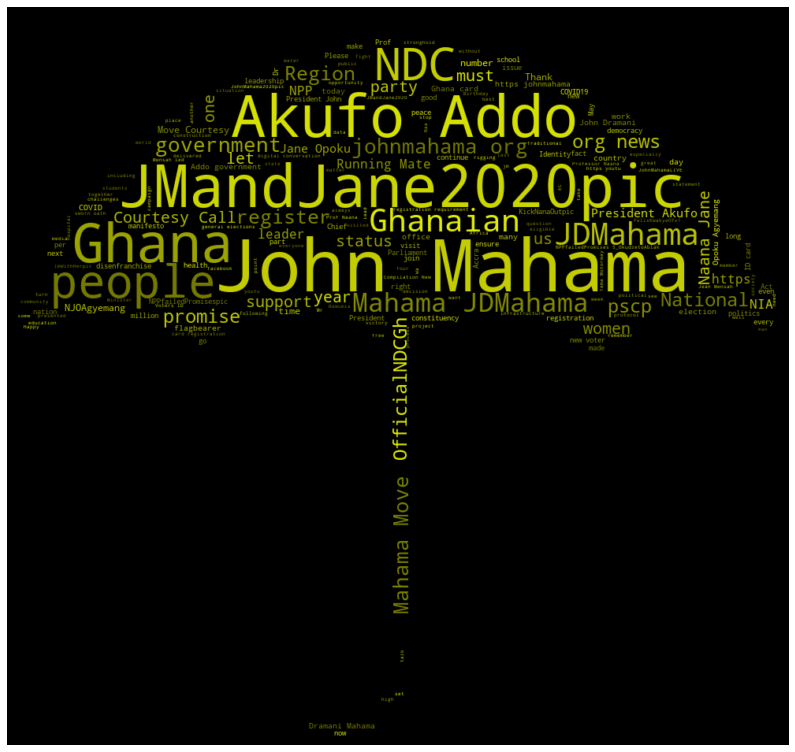
This huge variation could be to several reasons and the one that stands out most could be the different clout NJOAgyemang brings unto the ticket of the NDC. The credibility of the opposition party was gravely injured in the buildup to the last held 2016 general elections and the Ghanaian populace booted them out of office resulting in a huge landslide victory for the then opposition party, now incumbent.

There has never been a female running mate of any major political party until in August 2020, when the NDC announced they would be going into general elections with one. This brought a bit of credibility to the ticket of the NDC and proved to be a decision many Ghanaian voters' side with.

So, per the above graph 2, it is quite clear that the account of NJOAgyemang has been very active and buoyant. This does not really represent how the elections result might go but it is obviously a good place to start

***3.2 Word Clouds***

***fig A***



*fig B*

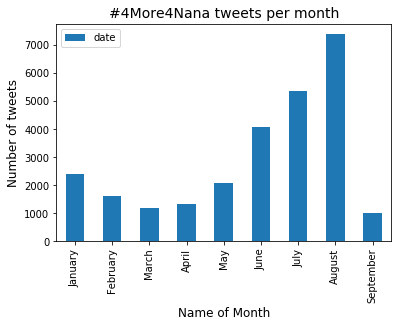
Next, we looked at the diction the leaders of both political parties. Word clouds are a great way to easily visualize how certain keywords stand out in a large collection of text. We used the WordCloud Python library to generate the images by feeding in the raw tweet data

Fig A represented the choice of words, the incumbent **H.E Nana Akuffo Addo** uses a lot. Words like **trust, deliver, promises** goes a long way to reverberate his accomplishments of policies or pact he made with the Ghanaian citizenry before assumption of office

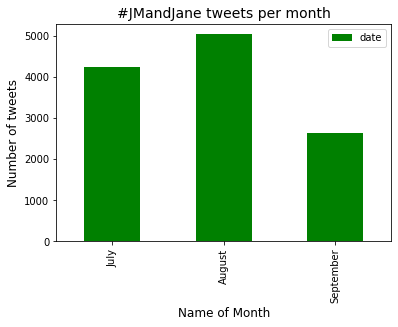
Fig B represents the choice of words used by the former president **John Dramani Mahama** most often. And it can be observed that he keeps echoing the name of running mate a lot, **JMandJane.**

The former president understands the huge credibility his choice of running mate brings to his ticket and he is also hammering that as often as possible and hopes to reconnect with the masses whose vote he lost in the last general elections

***3.3 Hashtags***

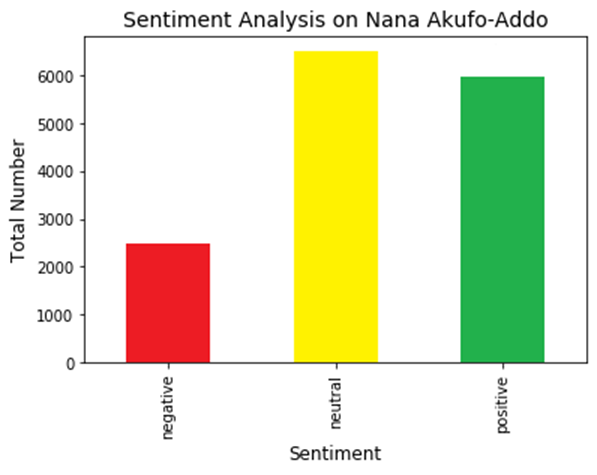
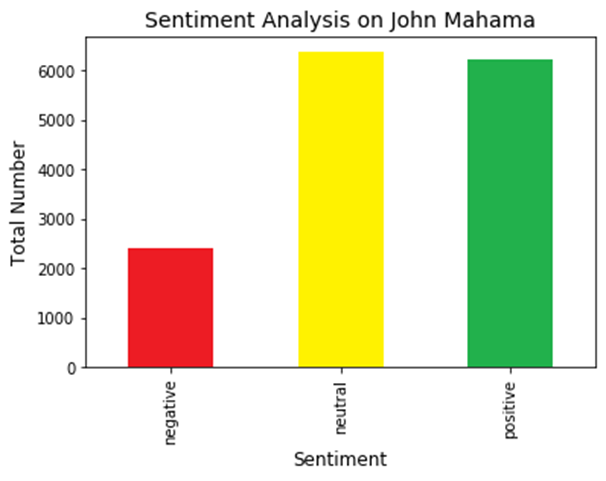


Graph 3

 Graph 4

Hashtags are used to represent trending topics on twitter. Popular hashtags used by the NPP since the month of January to September is 4more4Nana and for the NDC fraternity, JMandJane. Graph 3 shows the number of 4more4Nana from the month of January to September. From graph 3, it can be observed that 4more4Nana hashtags was a little over 2000 in January. These hashtags experienced rapid dip in numbers during the start of the ‘covid’ months and after 3 weeks constant lockdown, the citizenry gradually had their groove back and tweeting picked up

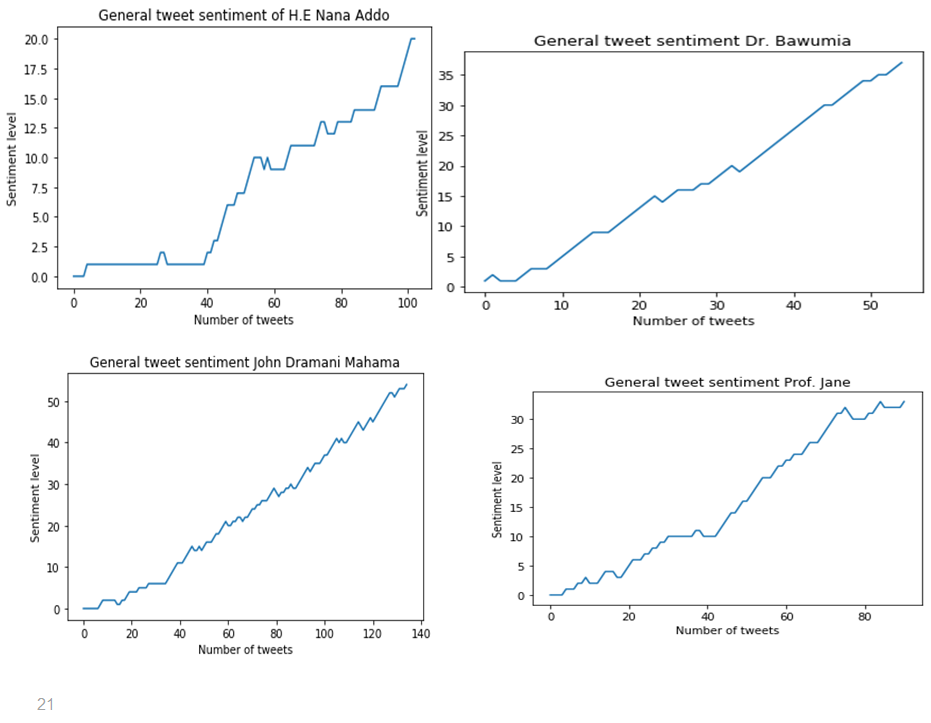
***3.4 Sentiment Analysis***



From the above graphs, it can be observed that both party leaders have an almost equal positive tweets, neutral tweets and negative tweets

Thus, the ghanaian culture goes hard on extremists or people who do not tackle issues and genuinely addresss them hence political leaders are very conscious of this culture and tweets accordingly.

***3.5 Sentiment level per number of tweets***



Sentiment levels of John Dramani Mahama and Prof. Jane is very high in their tweets as compared to their direct opponents, Nana Akuffo Addo and Bawumia respectively. This could be due to the natural compassionate natures of both Mahama and his running mate. Also, they tend to relay feelings of the masses because, they are in opposition and can sympathize easily with the citizens

***4.0 Conclusion***

* Increased political activity
* Political twitter has come to replace traditional media
* Followers of political leaders on twitter cannot be conclusively use to estimate the number of supporters a party has
* Tweets are considered official statements or opinions on matters of the state