OPINION MINING SYSTEM FOR PREDICTING ELECTION OUTCOME USING ASPECT-BASED SENTIMENT ANALYSIS

SANI, Y. M[1]., ETUK, S. O[2]., ADAMU, M[3]., & ADAMU, A. A.[4]
[1,2,3] Department of Information and Media Technology, [4]Department of Computer Science, Federal University of Technology, Minna, Niger State, Nigeria
E-mail: yahayasani@futminna.edu.ng[1], abiolastella@futminna.edu.ng[2]
bejian2004@yahoo.com[3] adamu.auwalu@st.futminna.edu.ng[4]
Phone no: +234-806-591-8550, +234-703-567-4352, +234-807-977-5565

Abstract
The fascination to predict the future is one of the intent and desires that humanity aim to achieve. These days, general public react to things and share opinions by the means of social media. This behaviour is on daily increase as people tend to divulge feelings and thoughts over the internet but sometimes this information is false and sometimes not useful. This study contributes to the emerging research on sentiment analysis using aspect-based model of social media contents related to a certain political events and trending topics in the political landscape. The system was implemented using Hypertext Preprocessor (PHP), Hypertext Markup Language (HTML), Cascading Style Sheet (CSS). This study focused on two (2) main political parties in Nigeria: All Progressives Congress (APC) and Peoples Democratic Party (PDP). Peoples' sentimental opinions were gathered (both positive and negative) and analyzed. It tested the difference in the polarity of the sentiments. The result shows that there is a difference in positive and negative sentiments and especially a significant difference in the negative sentiments between the political parties and how it affects the election outcome.

Keywords: Opinion Mining, Sentiment Analysis, Election Prediction, Aspect-Based Model

Introduction

Opinion mining is referred to as a type of natural language processing that can be used in tracking feelings and attitudes of the public about a particular topic (politics, entertainment or sports), services and products. In mining opinion, one of the growing fields of study is Sentiment Analysis. It helps in collecting information on people's opinions of a particular entity which is aimed at analyzing the positive and negative aspects of public opinions towards the particular entity (Bing Liu, 2012). Over time, social media as a key communication medium over the internet has being of great importance as majority of the world's population use the internet to publish their opinions on different topics and discuss latest trends (Brahmbhatt & Risha, 2017). One of such trending topics is politics and electoral process. If an aspiring candidate can predict earlier his success rate, it will help him in decision making but this seem to be a challenge. This work is aimed at helping candidates to anticipate their success by analyzing the public opinions about him (both positive and negative) and also provides an opportunity to know about the public demand and national will about their preferred subject (Amin, Seyyed, & Laya, 2013). According to Erik, Schuller, Yunquing and Catherine (2013), collecting opinions on the web will still require processing at the content/ syntactic level, filtering out un-opinionated user-generated content (subjectivity detection) and evaluating the trustworthiness of the opinion and its source. Various researches have been carried out on sentiment analysis in recent years. Since the tremendous improvement of web technologies, and the spectacular user engagements, enormous users' information is generated on various online platforms. These large amounts of information are used as data sources in applications based on sentiment analysis and opinion mining (Seureanu & Christian, 2012). The aim of this study is to use opinion mining system using Aspect-Based Sentiment Analysis to predict election outcome of the 2019 Local

Government Council Election in Chanchanga Local Government Area, Minna, Niger State, Nigeria using aspect-based sentiment analysis.

Literature Review

Aspect-based sentiment analysis (ABSA) is a text analysis technique that breakdown text in to aspects (attributes or components of a product or service) and allocates each one a sentiment level. It is all about listening to peoples view, understanding their voice, analyzing their feedback and learning more about their experience as well as their expectations for future or upcoming events. The big difference between sentiment analysis and aspect-based sentiment analysis is that the former only detects the sentiments of an overall text, while the later analyzes each text to identify various aspects and determine the corresponding sentiment for each one. In order words instead of classifying the sentiment of a text in to positive or negative, aspect-based analysis allows us to associate specific sentiments with different aspect of an event. The results are more detailed, interesting and accurate, because aspect-based sentiment analysis looks more closely at the information behind a text.

Below is the breakdown of what aspect-based analysis can extract:

(i). Sentiments: Positive or negative opinions about a particular aspect.
(ii). Aspect: Thing or topic that is being talked about.

Importance of Aspect-Based Sentiment Analysis

People are vocal; they love leaving feedback either good or bad making them valuable resources for business or to predict an event. Aspect-Based Sentiment Analysis takes all these feedbacks and automatically structures it so that companies and entities can interpret entries from customers and gain meaningful insights.

Reviews of Related Work

Jaidka, Ahmed, Skoric and Hilbert (2018) and (Elvyna & Yustinus, 2016) collected data from twitter post to predict election outcome. While the former used Naïve Bayes and standard lexicon approach for positive and negative sentiment, the latter used Naive Bayesian predictive model for each candidate and comparing the prediction with RealClearPolitics.com. They both concluded that mining Twitter for predicting election outcomes is a promising technique but they did not evaluate their findings.

Amin, Seyyed, and Laya (2013), proposed an opinion mining to predict the 11th presidential election of Iran and anticipate their participation. In their work, they used the K-Nearest Neighbors algorithm, classification tree and Naïve Bayes. They also used Orange data mining tool which proved to be efficient. Although their research was successful, KNN has shown to perform highly accurately in anticipation and classification compared with classification tree and Naïve Bayes. Their work was however, a proposal as it was not implemented.

Kolagani, Negahban and Witt (2017) analyzed public sentiments associated with the candidates in the 2016 United States presidential election. In their research, they collected over 200,000 tweets for the two major presidential candidates via hash tags. They designed a customized dictionary-based algorithm which they used to classify the tweets in terms of positive and negative as well as eight other types of sentiment (anticipation, anger, fear, joy, sadness, surprise, disgust, trust). The result of their study showed significant difference among the candidates in terms of surprise, joy, fear, disgust, trust and also positive sentiments, while the difference in the rest of the sentiments were insignificant. This system was only designed as it was not implemented.

MacDonald and Mao (2016) conducted a research using internet big data to predict United Kingdom' 2015 general election. They adopted text mining techniques, Big Data analytics and econometric models for sentiment classification. While their research gave a promising result, it was not possible to perform a direct seat prediction due to lack of access to data for prediction.

Hamling and Agrawal (2016) carried out a research to predict the 2016 United States presidential election using data they collect from tweeter. They developed an algorithm for sentiment classification and also developed a program to collect tweeter data. Though their work was time consuming and the algorithm they employed did not give promising result. From the above literature review, there is no research that adopted aspect-based sentiment analysis to predict the outcome of an election. This is the gap that this research aims to bridge.

Research Methodology

Dataset for Aspect-Based Model

From the dataset leads to model construction, the classifier model used is aspect level to classify texts as class labels: negative and positive. Each text provided in the dataset is used by the comment system to identify text polarity provided by the user. Table I below shows list of datasets used.

Table I: Dataset for Feature Extraction

Data Set (32)	
Negative words (16)	Positive words (16)
Bad, corrupt, thief, looters, crazy, hate, fail selfish, failed, lose, confuse, failure, incompetent, loot, wicked, poor, worst, corruption, looting, fool, negative, thieves	Care, good, great, integrity, plans, wailers interest, chance, ideas, ideology, win, succeed, competent, love, best, positive,

Proposed System Architecture

Users feedback inputted into the comment system is pre-processed by converting the comments into lower case. Then the keywords in the system database are extracted from the comments with the given attributes provided in the database. The polarity of the users' comment is then classified as positive or negative. Figure 1 below shows the system architecture of the proposed system.

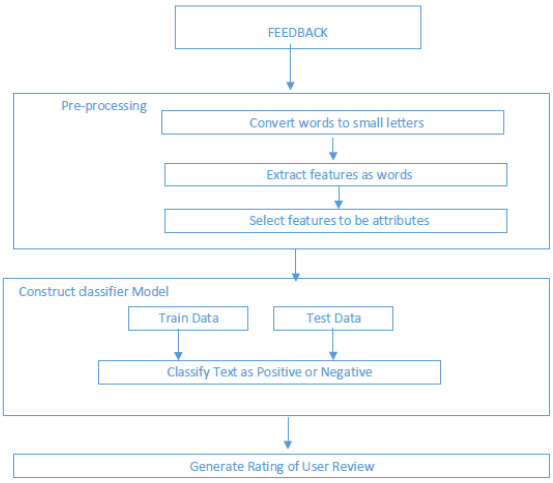

Figure 1: Proposed System Architecture

The classification is done to distinguishes data classes which consists of an attribute of positive and negative sentiments and the concepts involves the following steps:

(i). Learning Step (Training Phase): The construction of the model (Aspect level) with the above algorithm are used to build a classifier by making the model learn using the training set available. The model has to be trained for the prediction of accurate results.

(ii). Classification Step (Test Data Phase): This phase is used to predict class labels and testing the constructed model on test data and hence estimate the accuracy of the result.

System Flowchart Diagram

This is a visual representation that shows the sequence of steps and decisions needed to perform the process. The flowchart for the proposed opinion mining system using sentiment analysis is shown in Figure 2.

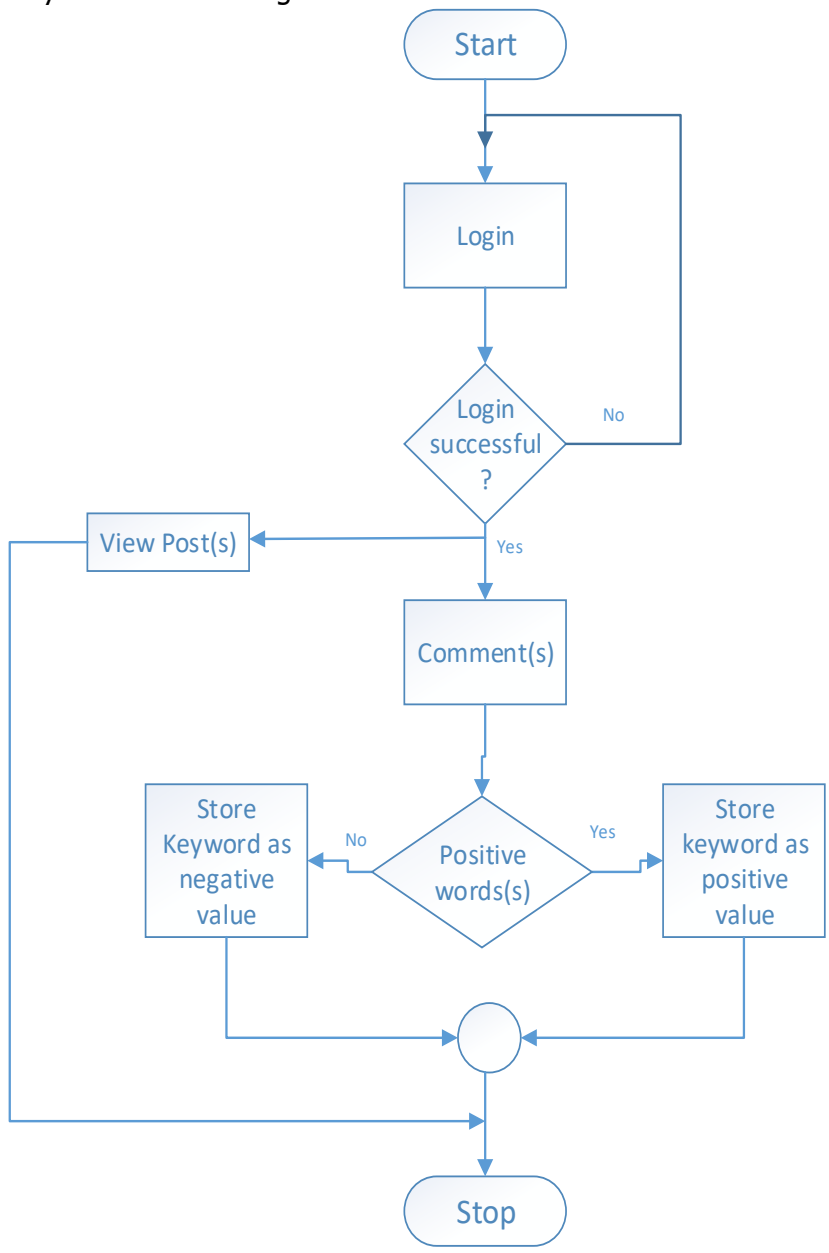

Figure 2: Flowchart of the Proposed System

System Use Case Diagram

The use-case diagram for the system depicts activities carried out by the actors elaborately. It shows what activities can be performed by the actor based on the system operations. The figure below shows how the user interacts with the system from user input to the system down to information querying by the respective actors. The diagram is shown in Figure 3.

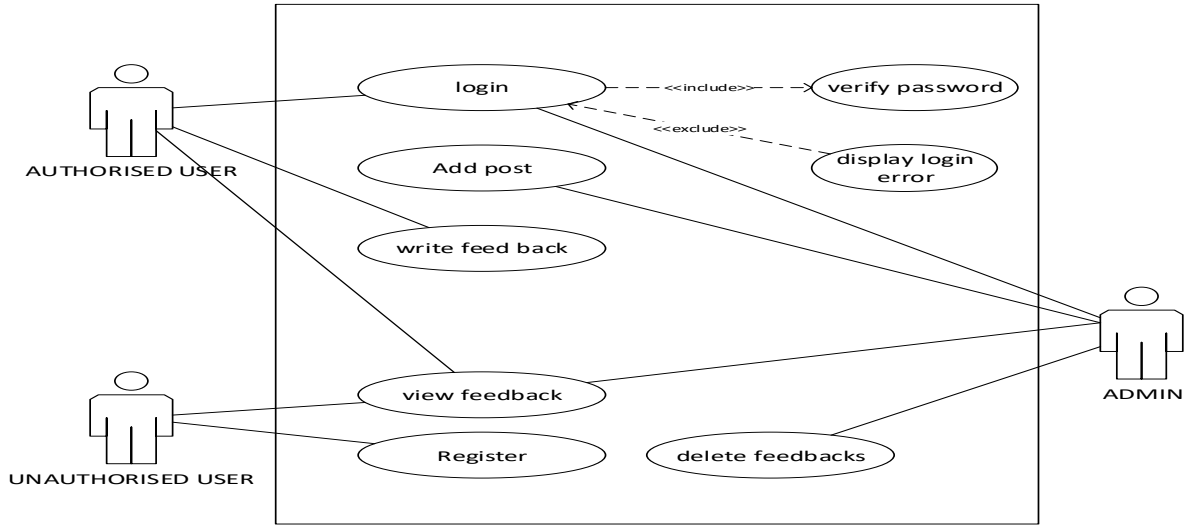

Figure 3: System Use Case Diagram

System Implementation

Program Setup: The system is setup by initializing PHP artisan serve using command prompt and running the executable files published from the PHP, HTML and Java codes were used to implement the system and runs on a web browser. This is done alongside the installation of other required software such as composer setup to activate the PHP artisan server, Mysql for database and XAMPP server.

Login Page: Authorized registered users can login to the system to perform the rights given to them through their user email and passwords. The login page is shown in Figure 4.

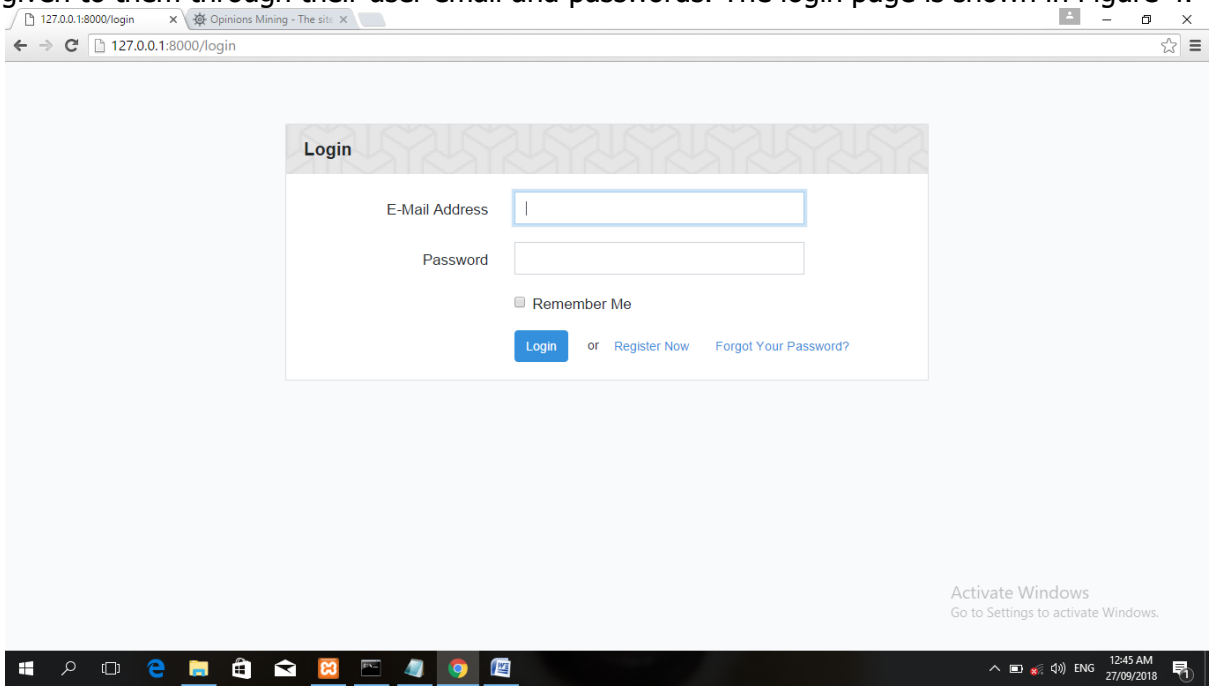

Figure 4: Login Page

Admin Dashboard

The admin dashboard authorized admin to have users view, control users access and posts. The admin dashboard is accessible through a valid admin email and password. The Figure 5 shows the admin dashboard.

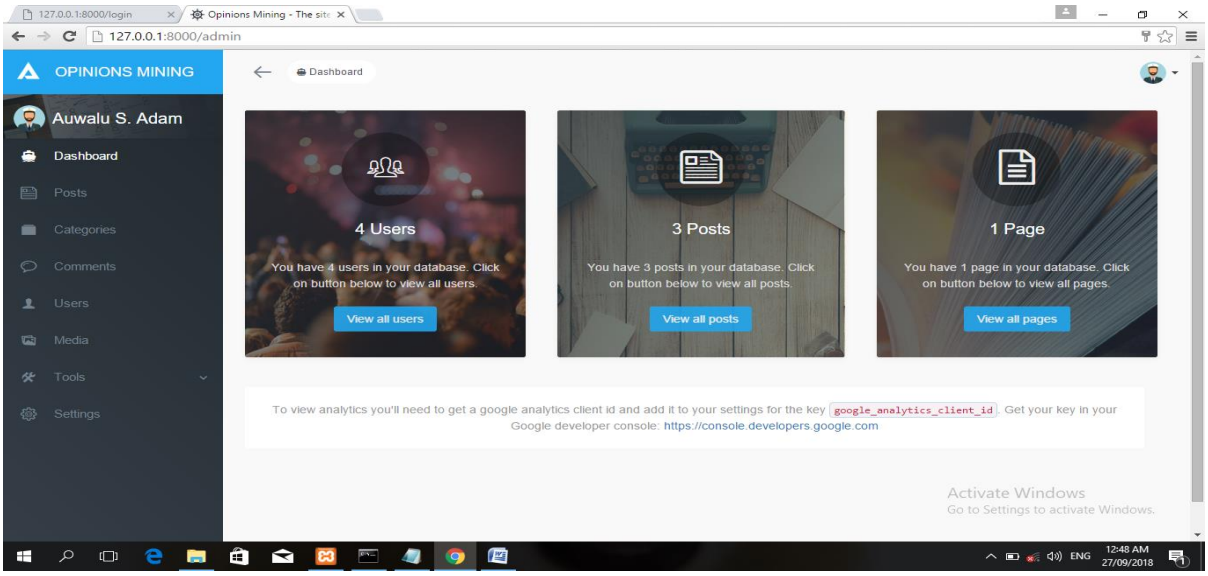

Figure 5: Admin Dashboard

Add Review

The add review section is where registered users input their feedbacks to the comment system. The comment system has been provided with keywords to check the polarity of users' review. The add review system is shown in Figure 6.

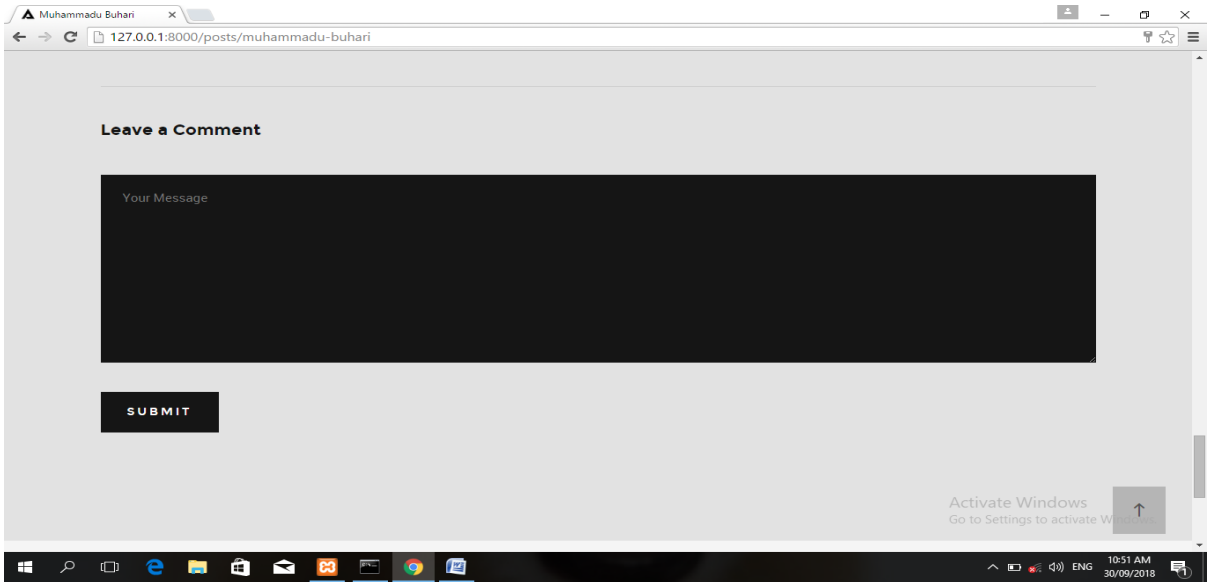

Figure 6: Add Review

Comparison of Aspect Extraction for both Positive and Negative Feedbacks

Here, the positive and negative feedbacks collected from both parties are compared for real analysis. This is shown in Figures 7 and 8.

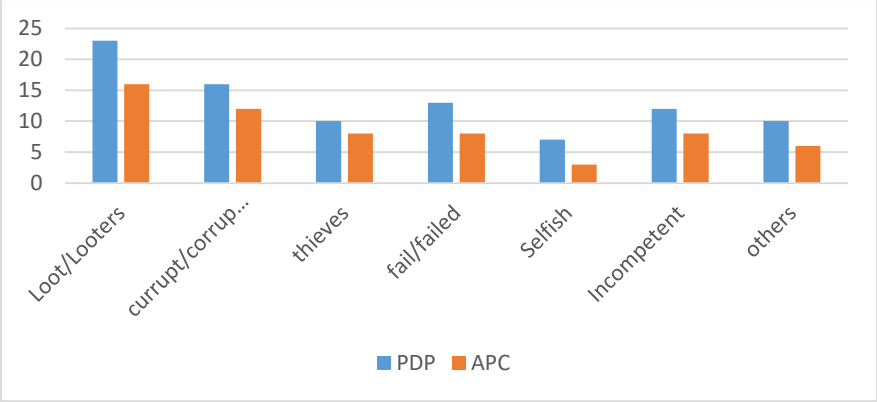

Figure 7: Aspect Extraction for Negative Feedback

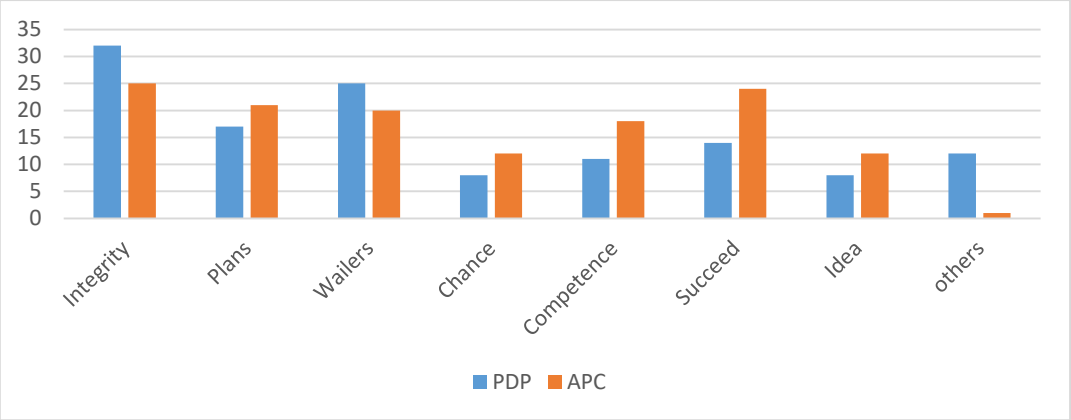

Figure 8: Aspect Extraction for Positive Feedback

Results and Discussion

A total of 260 intending voters for the 2019 Local Government Council Election in Chanchanga Local Government Area, Minna, Niger State were randomly selected from 13 polling units in the Local Government Area to test run the system. 127 intending voters gave positive sentiments to the PDP party which is 48.85% of the total voters while 133 voters were attributed to the APC party which gives 51.15% of the total number of voters for election success chances. In a similar fashion, as shown in Table III, out of 162 voters who posted their opinion (negative sentiment), 91 voters posted against PDP which is 56.17% of the whole intending voters while 71 voters posted against APC which is 43.83%. The difference in the polarity of the sentiments in terms of positive and negative sentiments in general is found to have a significant difference in negative sentiments between the candidates, while the positive sentiments in general did not have a significant difference. The illustration is shown in Table II and III as well as Figure 9 and 10.

Table 2: Percentage Result of Positive Feedbacks for both Parties

Positive Feedbacks	Frequency (Word Count)	Percentage
APC	133	51.15%
PDP	127	48.85%
Total	260	100%

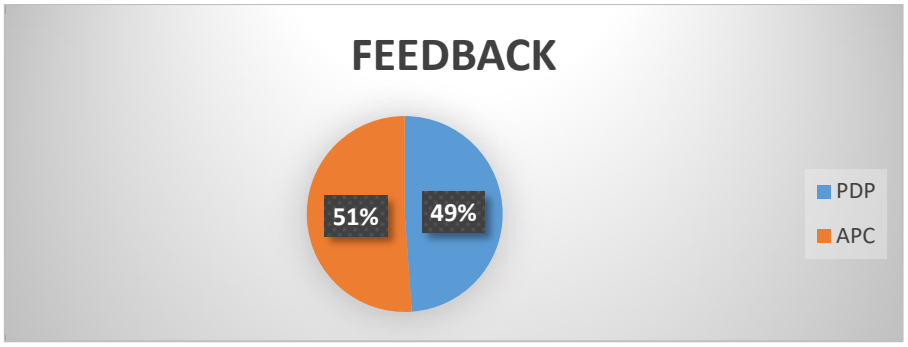

Figure 9: Representation of Result of Positive Feedbacks for both Parties

Table 3: Percentage Result of Negative Feedbacks for both Parties

Negative feedbacks	Frequency (Word Count)	Percentage
PDP	71	43.83%
APC	91	56.17%
Total	162	100%

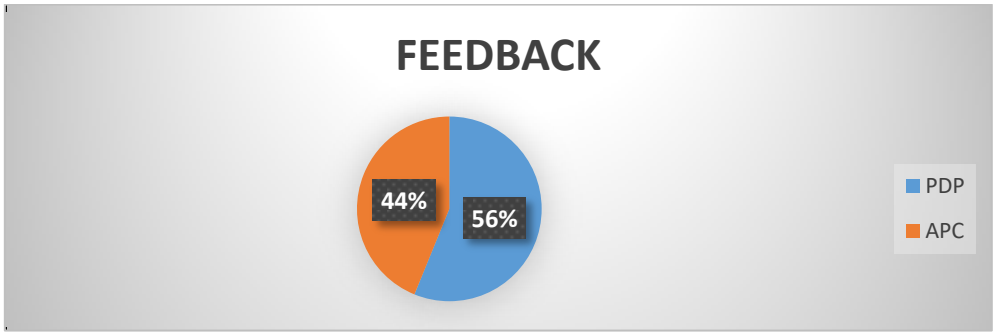

Figure 10: Representation of Result of Negative Feedbacks for both Parties

Conclusion

An opinion mining system to predict election outcome using Aspect-Based Sentiment Analysis was implemented. The system was implemented using hypertext pre-processor language, XAMPP, MySql database. The developed system has proven to be efficient and effective in predicting the outcome of election since the APC is most likely tipped to win the election as predicted.

References

Amin, B. S., Seyyed, R. K., & Laya, E. (2013). Participation anticipating in elections using data mining methods. *International Journal on Cybernetics and Informatics*, 2, 47-60.

Bing, L. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures of Human Language Technologies*, pp 6-167.

Brahmbhatt, A., & Risha, T. (2017). Opinion mining to predict election results. *International Journal for Technological Research in Engineering*, 7(4), 2347-4718.

Elvyna, T., & Yustinus, E. S. (2016). Bayesian Twitter-Based Prediction on 2016 U.S Presidential Election, pp 12-39.

Erik, C., Schuller, B., & Yunqing, X. H. (2013). Catherine new avenues in opinion mining and sentiment analysis. *Knowledge Based Approaches to Concept Level Sentiment Analysis*, 15-21.

Hamling, T., & Agrawal, A. (2016). Sentiment analysis of tweets to gain insights into the 2016 U.S election. *Columbia Undergraduate Science Journal,* pp 34-42.

Jaidka, K., Ahmed, S., Skoric, M., & Hilbert, M. (2018). Predicting elections from social media: A three-country, three-method comparative study. *Asian Journal of Communication*, 1-22.

Kolagani, S., Negahban, A., & Witt, C. (2017). Identifying trending sentiments in the 2016 u.s presidential election: a case study of twitter analytics. *Issues in Information Systems*, 2(18), 80-86.

MacDonald, R., & Mao, X. (2016). Forecasting the 2015 general election with internet big data: An application of the TRUST framework, pp 5-40.

Mishra, N., & Jha, C. K. (2012). Classification of opinion mining techniques. *International Journal of Computer Applications,* 56(13), 0975-8887.

Smeureanu, I., & Cristian, B. (2012). applying supervised opinion mining techniques on online users' reviews. *Information Economica*, 16, 81-91.