International Journal of Sciences, Engineering & Environmental Technology (IJOSEET), 6(3): 24 – 30, 2021
ISSN 0794-9650
www.repcomseet.com

Sentiment Analysis on Influence of Technology on Society

Lawal O. Moshood[1*], Lawal T. Aderonke[2], Ogunleye T. Omotayo[3], Akporherhe A. Oniovosa [4]
[1,2,3,4] Computer Science Department, Federal Polytechnic Ede, Nigeria

*Corresponding Author: *Lawal M.O. (olawal70@gmail.com).*

Abstract: This research work is based on mining and analyses opinions using natural language processing. Sentiment analyses were used to verify the impact of technology on communities, and it majorly focuses on obtaining raw text data to check how much technology impacts society using Twitter handles and streaming raw text data from Twitter API. Naïve Bayes Classifier was used in a Python algorithm to analyze the seven thousand streamed raw twitter data from the Twitter API using two different Twitter handle (@YouthForTech, @TechRepublic) to generate opinion (positive, negative, and neutral opinion respectively). The process was implemented using the Naive Bayes classifier in Python algorithm and it employs Tweepy and NLTK library. The algorithm automatically creates a "JSON" file that contains the dataset. And analyses were performed on the dataset to generate tokenized format of the data. Hence, the result was used to generate a pictorial presentation of the result in the form of charts that show the percentages of the positive, negative and neutral opinion of filter/search parameters that were analyzed by the sentiment analyzer. The results also show that the positive effects of technology are higher than both the negative and neutral values obtained from the analyses.

Keywords: Sentiment Analysis, Naive Bayes Classifier, Technology, Society, Twitter API, Streaming.

1.0 Introduction

The growth of technology has risen to the level, that its effects are now felt in our day-to-day routines. And the advancement in technology makes it seem like it now controls our lives. The availability of technology and its uses is broadly encouraged throughout diverse communities. Technology makes day-to-day routines easy; it also creates some social problems such as reduced social cultures and manners. The use of technology has more effect on the daily lives of the younger generation and on most occasions, they seem to be preoccupied with the use of different technological devices like smartphones, tablets, game consoles, etc. Hence, technology use has increased tremendously and this has affected the daily lives of communities and the young generation. As technologies grow, their effects are felt through generations. The last few decades show that mobile devices, personal computers (PCs), and indispensably the internet as completely changed the interrelationship between communities and how instructions are given in schools. The application of technology in learning environments is a vital tool for the successful completion of education from primary to university educational level. The demand for technological clever personnel has been the major reason for the growth in innovation and technology in schools today. Educators must adapt to the millennial advances in the technological wheel make sure that their tutee is also prepared for new innovations in technology. The outlook of society has been altered by technological advancements, classrooms are now technology-based and the use of chalkboard and letter writing will no longer be available. The demand for technologically savvy personnel has also increased the demand for 21st-century classrooms (Brian, 2013).
A society encircles any assembly of a category of persons living together in a community, which comprise some form of regulations, along with laws, roles, and an economy. Society often enclose factors of agriculture in the economy, along with goods importation and exportations, such communities often include militaries, educational centers, technological hubs, etc., and they grow into technologically advanced countries (Wardynski, 2019).

Fact-finding techniques involve a huge amount of text data. Analyzing such text data usually requires a large number of computer resources. However, it is estimated that eighty percent (80%) of the world data are unstructured meaning that the data are unorganized. Such huge volumes of text data come from emails, social media, conversations, surveys, articles, documents and so much more. However, analyzing such text data, sort through can be very difficult not to mention time-consuming and expensive.

Sentiment analysis also referred to as opinion mining is the process of using computing to understand whether a piece of raw text data is positive, neutral, or negative. Opinions are obtained from the behavior of a speaker. Sentiment analysis is a tool born from Natural Language Processing (NLP) and the main ideas are to detect subjectivity in text, extract opinion and classify all the opinions into positive, neutral, and negative sentiments. It is also referred to as the point of view of people, attitudes, thoughts, feelings, and way of thinking towards services and their attributes. Opinion mining of text data is analyzed in the form of true or false, the polarity of human sentiment expression, the polarity of the outcome, survey value (agree or disagree), good or bad, support or opposed (Alessia, 2015).

2.0 Literature Review

Lailah (2000), in his study describes the role of communications in information technology and stresses how technology is affecting the interpersonal relationship between families. It also focuses on the negative effects of technology on our society. The effects have disintegrated the moral and value systems in most family interactions. This paper also proposes that bring up the young generation should be done in accordance with the moral systems used within their identity.

Oblinger *et al* (2001), refers to the net generations as children growing up within technological communities, they are also called digital natives, millennials, net generations, and "grasshopper minds". The idea of millennials is generally referred to as children born between the years 2000 above. The net generations refer to children who grew up in an advanced country with rich technology, introduced to technology at an early stage, and are constantly using technology heavily.

Al Hawsawi, A. *et al* (2006) opined that the high rate of spread in modern technology has drawn major concerns in its uses and its gloomy impacts. The absence of proper guidance as increases the level of vulnerability in certain groups of the young generation and its negative effects have become life-threatening. However, education is an integral part of life, and modern technologies are far too important to education and not just simple addition.

Aga, (2009), pronounces that it is safe to say that technology is a principal part of the younger generation's lives. They can obtain rudimentary prowess of technology use very easily, speedily and they can creatively use the technology for various uses. Children define the use of technology as a form of a game, while adults think of technology as a functional tool for achieving positive things. They also consider that technology may be a problem if control and regulation are not imposed on its usage. This paper also discusses the apprehension about the peril of technology regarding health and uncensored use of websites over the internet.

Alessia et al, (2015), discuss various approaches to sentiment analysis, the paper describes the approaches with respect to advantages or limitations. Approaches such as machine learning, lexicon-based and hybrid were discussed. The techniques employed in each approach were stated and the advantage of each of these methods was also highlighted. Analysis done in this study using the different tools mention can be applied to different fields such as; finance, politics, societal differences, agriculture, etc.

3.0 Aim of the Research Work

This research work is aimed at developing an algorithm using Python programming language to mine Twitter data of selected information technology Twitter handles (such as @youthForTech, @Engadget, @TechRepublic, *@ZDNet, @CNET, @SmartPlanet,r @CBSi, etc.*) through the Twitter API. And extract opinions and classify all the opinions and their sentiments from the mined Twitter data using rule natural language processing.

3.1 Methodology

The aim was accomplished through the following approaches:

i. Twitter data mined in JSON format was parsed through the python programing language to generate an opinion of tweets on the handle;
ii. Implementing Naive Bayes classifier using Python, Tweepy, and NLTK library.
iii. The result of (i) is used to generate positive, negative, and neutral values and also create charts of the percentages of the opinions generated.

4.0 Twitter Sentiment Analysis

This paper presents an approach for classifying the sentiment of Twitter messages or tweets; these messages are classified as positive or negative with respect to a sentence. This was accomplished by mining tweets using Twitter search API and subsequently processing them for analysis. Moreover, this research work uses Distant Supervision, in which the training data consists of tweets with emoticons. Furthermore, an examination of the effectiveness of three machine-learning techniques was used to provide a positive or negative sentiment on a tweet corpus. The effectiveness of the algorithm was tested using the Naive Bayes classifier, Maximum Entropy classifier, and Decision Tree classifier. The research shows that the accuracy of those algorithms is above 60% when trained with emoticon data (Andrew L. 2011).

4.1 Tweeter Text Analysis Using Naive Bayes Classifier

The Naıve Bayes classifier (NBC) is the most commonly used classifier. It uses a classification model to compute the tail probability of a class, based on the arrangement of the words in the document. Walaa 2014 uses a feature extraction technique on the bag of words which ignores the position of words in a document. He also anticipates the probability that a given feature set belongs to a particular label using the Bayes Theorem. The formula below is used to describe the Bayes Theorem.

$$P(L \backslash F) = P(L) * P(F/L)/P(F) \quad (1)$$

i. P(L) is the probability that a random feature (F) sets a label (L) and the feature is also being classified as a label.
ii. P(F) is the probability that a given feature set has happened. The Naïve theorem assumes that all states in the features are independent and the equation is rewritten below;

$$P(L \backslash F) = P(L) – P(F1 \backslash L) * \ldots * P(Fn \backslash L) / P(F) \quad (2)$$

Building a sentiment analyzer, understanding the right methods and tools are very essential. One of such tools is machine learning which is used to create different methods of classification. Data training requires classifiers and supervised learning requires hand-labeled training. Discussions on Twitter are usually diverse and it would be difficult to manually collect and label all the data to train a sentiment classifier for all the tweets. A method that could be employed is the distant supervision that allows training data from tweets with an emoticon. Emoticons are noisy labels in tweets and emoticons like **"☺"** show positive sentiment and **"☹"** portray negative sentiment. Streaming a large number of tweets with emoticons from a Twitter handle or account has been made available through the Twitter API. And this is a major advantage over the hours it will take to hand-label training data Anssi Klapuri, (2007).

4.2 Training the Classifier

Training of data involves using an algorithm to forecast the output of a designed model in machine learning. A classifier such as Naïve Bayes can be used to creating a sentiment analyzer.

The naïve Bayes method is used in the Natural Language Toolkit (NLTK) library to train and classify data. Hence, with the NBC and NLTK, we have a training set, so all that is needed to do is instantiate a classifier and classify test tweets. Furthermore, the research work creates three files for;

i. Streaming the data set from Twitter search API which will create datasets
ii. The algorithm automatically creates a .json file that will contain the dataset.
iii. And thirdly, execute Naïve Bayes algorithms which will run the classifier and request for a prompt to enter a sentence (simulating a tweet) and this will finally show if the sentences have a positive or negative and neutral feeling.

Analyzing tweeter data involves breaking down unstructured raw text data, preprocessing it, creating a normalized format, and generating a statistical analysis of the data. The normalized data is tokenized to create tokens, words, phrases, and symbols from the tweets that were downloaded from the Twitter API. Because twitter contains information such (URL, Emoticon, # tags, etc.) is the reason why the tokenization takes place to remove noise within the data. Consequently, preprocessing steps such as natural language toolkit (NLTK) uses a tokenizer class (tweetTokenizer) to tokenize the tweets, and other classes such as stop word removal (such as punctuations, lol, RT, via, etc.), nltk.corpus.stopwords, and case normalization to break down the tweets into a readable format (Walaa, 2014).

4.3 Fetching Tweets from Twitter Handle Time Line

Fetching of data from Twitter through the application programming interface (API) creates streams of raw text data from a specified Twitter account using a python algorithm. The raw data is downloaded to a personal computer in JSON (standard for file format and data interchange for data objects) format. Fetching raw data from Twitter can be done in different ways; using the REST API and Streaming API.

i. REST API: this involves going back in time to fetch existing tweets. These are tweets that are already published and are made available for search.
ii. Streaming API: This approach looks into the future. It involves keeping the HTTP connection to a Twitter account open and retrieving all the tweets that match the filter/keyword search criteria as they are being published in real-time. It is the preferred method of downloading a huge amount of tweets. This method can be time-consuming as fetching involves waiting for the tweets to be published.

This research work makes use of streaming API to generate its raw data and over 7000 tweets are downloaded from two different Twitter handle (@youthForTech, @TechRepublic). Below is a simple illustration of how tweets are retrieved from the Twitter API.

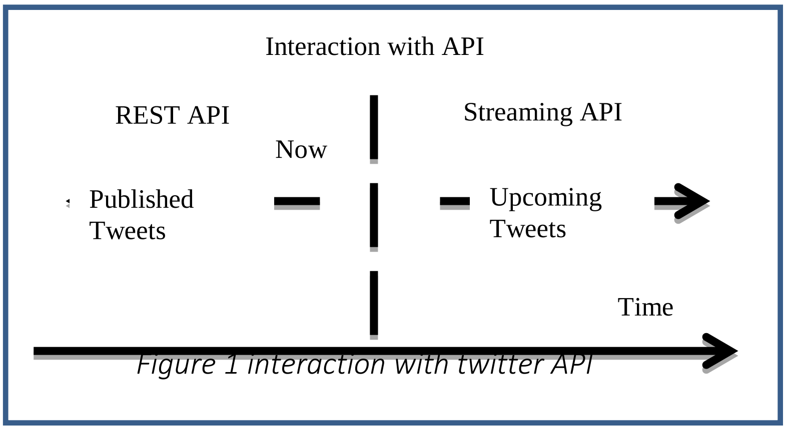

Figure 1 interaction with twitter API

4.3.1 Result of Sentiment Analysis to Generate Pie Chart

Using the algorithm developed, two Twitter handles were streamed and download. Sentiment analyses were performed on the tweets using Naive Bayes Classifier (NBC) and it was discovered that the first handle (YouthForTech) has Positive tweets percentage: 77%, Negative tweets percentage: 5%, Neutral tweets percentage: 18% have used young generation and technology as the sentiment analysis filter. Hence, the same process was used to perform analysis on the @techRepublic handle to generate; Positive tweets percentage: 44%, Negative tweets percentage: 22%, Neutral tweets percentage: 33% have used the effect of technology as the sentiment analysis filter.

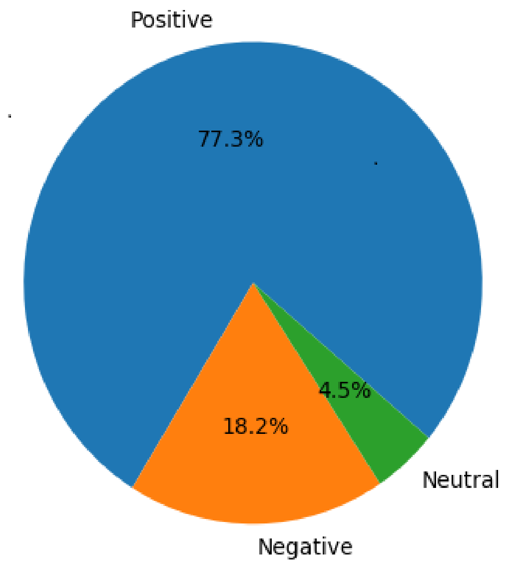

Figure (2) YouthForTechnology Pie chart

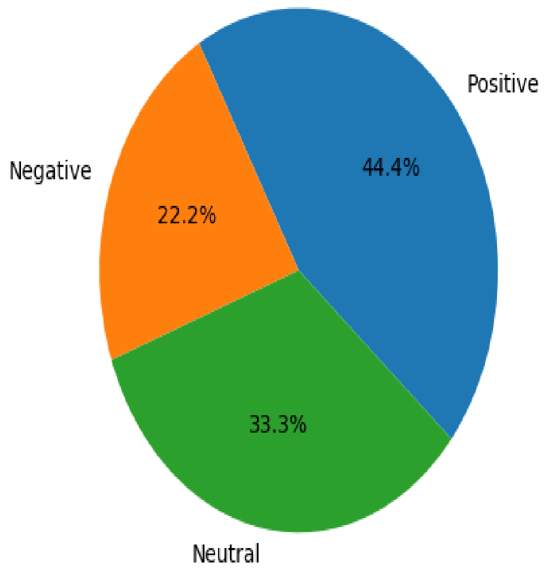

Figure (3) TechRepublic Pie chart

Figures (2) and (3) above shows the graphical representations of the result using pie chat. It shows the percentage of the positive, negative, and neutral opinions of how society perceives the use of technologies.

4.3.2 Result of Sentiment Analysis to Generate Bar Chart

Using the Naive Bayes Classifier (NBC), figure (4) below shows the bar chart for the sentiment analysis perform on the streamed raw data from the search filters (keywords) used to retrieve the streamed data. The raw data was stored using .json format and saved in the same location as the algorithm for ease of access by the algorithm.

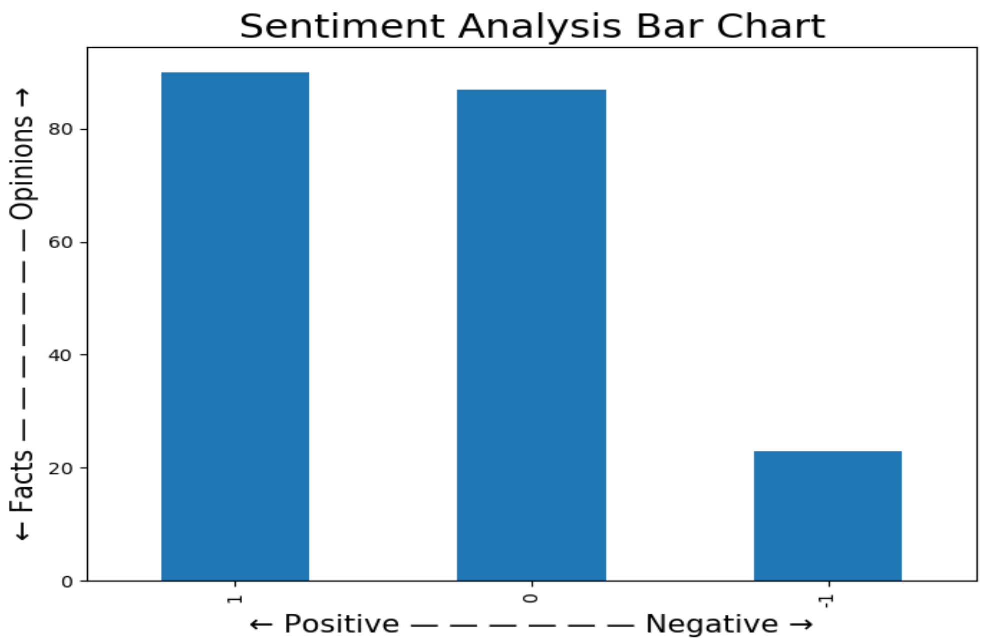

Figure (4) Sentiment Analysis Bar Chart

4.3.3 Positive Impact of technology on Society

There are probably more constructive effects on society than negative consequences. Form the sentiment analysis in Figures (2) and (3) and (4) above, using different parameters and search filters, from the sentiment analysis, it shows that 77.3%, 44.4%, and 80% positive opinions respectively. It also reveals that the beneficial effects of technology on society are very high. Such effects have made life easier for many, and have made resources, education, and tools available to many resulting in living a better life. Such impacts have also greatly influenced agriculture, communications, and businesses within societies worldwide.

4.3.4 Negative Impact of Technology on Society

Using sentiment analysis, the charts in figure (2) and (3) and (4) respectively, shows that using different parameters and search filters, the percentages of opinion from tweets show that 18.2%, 22.2%, and 20% negative opinions respectively. There has not been a single form of innovation that does not have its deficit. The results above show that even though a lot of people use technology for various activities in their day-to-day life, its use has also resulted in some form of hazard to human health, the environment, and possibly agricultural development. Technological advancement has also been a major cause of warfare between nations and a tremendous cause of global warming. The uses of technology also reduce natural resources in form of rare earth minerals and a major reason for the rise in population due to advanced medical facilities and equipment.

4.3.5 Neutral Opinion from Sentiment Analysis

Using the sentiment analysis in Figures (2) and (3) and (4) above, using different parameters and search filters, and it shows 4.5%, 33.3%, and 70% neutral opinions respectively. This result shows that some users of technology are not particularly attached to a specific technology. Their usage of technology is based on the need of the individual, persons, and young generations that falls within the neutral position at that specific point in time.

5.0 Conclusion

Th e use of technology has been a major reason for the development in advanced and developing nations. This research work uses an aspect of information technology to sample people's opinions on the acceptability of technology and how it has affected communities and society. While technology is considered to make life better, it is also an instrument of production for societies. Hence, the use Naïve Bayes classifier for sentiment analyses in natural language processing was used on Twitter streamed raw data employed to determine the opinion of the people using the specific Twitter handle. However, the raw data collected was streamed from the last seven days tweets from the Twitter handle and over 7000 tweets were collected in JSON format. The results show the opinion in the form of charts and a pictorials representation of the opinions in percentages. It also shows that there are more positive opinions about technology more than negative and neutral opinions. This research work shows that technology is a phenomenon that makes people's lives easier having examined the positive, negative, and neutral opinions of tweets from technology-based Twitter handles. Consequently, it is suggested that for further research work, the raw data collection should be done with the use of REST API to further broaden the scope of the search data to months of the backlog of tweets.

REFERENCES

Aga Al, (2009), The misuse of the Internet and mobile technology and their role in juvenile delinquency in the GCC,.

Al Hawsawi, A (2006), Impediments to the use of educational technologies in teaching mentally retarded students as perceived by teachers of Education in Riyadh..

Al Yousef, Sh, (2006), Modern technologies advantages and disadvantages of a study of the negative effects on the health of the individual.

Alessia D'Andrea et all, (2015), Approaches, Tools and Applications for Sentiment Analysis Implementation (available online : http://www.ijcaonline.org/research/volume125/number3/dandrea-2015-ijca-905866.pdf)

Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts, (2011). Learning Word Vectors for Sentiment Analysis.

Brian Sutton, (2013), The Effects of Technology in Society and Education,

Duggan M., Lenhart A., Lampe C., Ellison N. B., (2015). *Concerns about children, social media, and technology use*. http://www.pewinternet.org/2015/07/16/concerns-about-childrensocial-media-and-technology-use/ Presiden erişilmiştir.

Gökhan Dağhan, (2017), Views of Students about Technology, Effects of Technology on Daily Living and their Professional Preferences, TOJET: The Turkish Online Journal of Educational Technology – October 2017, volume 16 issue 4

Gregory Grefenstette , (2010), Use of semantics in real life applications. Publication at:https://www.researchgate.net/publication/221613857

Karehka Ramey, 2012, https://www.useoftechnology.com/technology-society-impact-technology-society/

Kathleen m., Ralph Wilburn, (2018), the impact of technology on business and society. Global Journal of Business Research.

Lailah, A, The role of mass media and information technology in the family disruption. p. 23, 2009.

Oblinger, D. G., & Oblinger, J. L. (2015). *Educating the Net Generation.* EDUCAUSE: Transforming Education through Information Technologies.

Pang Bo, Lee Lillian (2008), Opinion mining and sentiment analysis. Found Trends Inform Retrie.

Peter Wahler, Claus J. Tully, (1991), Young people's attitude to technology. European Journal of

Education.

Walaa Medhat, Ahmed Hassan, Hoda Korashy. Sentiment analysis algorithms and applications: A survey; 2014.

Wan X, (2009). Co-training for cross-lingual sentiment classification. in: Proceedings of the joint conference of the 47th annual meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing (ACL/IJCNLP'09); 2009. p. 235–43.

Wardynski, D. J. (2019), Technology and Society: How Technology Changed Our Lives https://www.brainspire.com/blog/technology-and-society-how-technology-changed-our-lives
https://www.brainspire.com/blog/technology-and-society-how-technology-changed-our-lives