Spektrum Industri
Vol. 21, No. 2, 2023, pp. 100-108
ISSN 1693-6590
http://journal3.uad.ac.id/index.php/spektrum

Emotion Detection for Health and Well-being in Short Messaging Systems

Emeka E. Ogbuju*

Department of Computer Science, Federal University Lokoja, Lokoja, Kogi State, Nigeria.
*Corresponding Author: emeka.ogbuju@fulokoja.edu.ng

ARTICLE INFO

Article history
Received July 15, 2023
Revised August 10, 2023
Accepted October 27, 2023

Keywords
Emotions detection;
Short message service;
Logistic regression;
Machine learning;
Text analysis

ABSTRACT

The exposure to unpleasant emotions or content in messages can lead to health complications, including high blood pressure and several heart-related disorders. Hence, the identification of unpleasant emotions in written content can serve as a beneficial instrument in addressing certain health-related issues. Emotions can be communicated through diverse modalities, including written text, spoken language, and facial gestures. The objective of this work is to create a Text-based emotion detection system that possesses the capability to accurately identify emotions within text messages. The use of message filtering mechanisms that detect and block content containing negative emotions can serve as a preventive measure to shield users from accessing messages that have the potential to adversely impact their well-being. Conversely, messages that convey positive or neutral emotions remain accessible for comprehension. In order to accomplish this objective, a combination of three machine learning algorithms, namely Naive Bayes, Support Vector Machine, and Logistic Regression, were employed, adhering to the CRISP-DM approach. The Logistic Regression technique achieved the greatest accuracy rate of 98.4% and was employed in the construction of the detection system. The Graphical User Interface (GUI) of the system was developed utilizing HTML and CSS, with the integration of diverse components to establish a comprehensive and operational interface for the user.

1. Introduction

The third Sustainable Development Goal (SDG 3) aims to achieve universal health coverage and promote well-being for all, with a specific target to eliminate the epidemics of communicable diseases, including AIDS, tuberculosis, malaria, and others, by 2030. It is line with this that this study seeks to explore technological ways to achieving this goal. Emotion is a principal part of human existence. It is an individual's internal state of being and an automatic physiological reaction to an object or situation, which may be linked to a physical state and sensory information (Ho and Cao, 2012). Krishnan (2017) defined emotion as a particular feeling that characterizes a state of mind, such as joy, anger, love, fear and so on. Emotion detection is an important field of research in human-computer interaction, due to the fact that many persons rely heavily on a computer to achieve certain task, and to convey their emotions to one another, which are generally achieved through verbal, non-verbal (includes expression through facial, gesture or body movement) and text-based channels. In recent times, the World Wide Web has emerged into a smart-mobile industry which allows humans to communicate with each other through various social media sites (such as Facebook, Twitter, and Instagram), video sharing sites (such as YouTube) or blog. Humans widely interact with computer via texts while multimodal human-computer interaction is also denoted to be appealing. It cannot be denied that the short message service (SMS) remains one of the earliest text-based mode of interaction, even before the advent of the various social media platforms. However, due to the influence of this new media on the society, most of the people like to share their opinion or express their emotions via any of the social media platforms. Hence, various studies on emotion detection and control have been carried out on the aforementioned social media outlets.

In line with the foregoing, it is plausible to state that much emphasis has not been placed on the SMS which is one of the common text-based modes of human interaction. Thus, efforts should be made to carry out a text-based emotion analysis on this mode of interaction. Emotion recognition, particularly text-based, remains a viable path for modern scholars and researchers, as its importance tilts towards humans need for communication and socialization. Emotion recognition (also known as: Artificial Emotional Intelligence) is an affective computing whereby the development of systems utilizes the process of detecting, interpreting, and predicting human emotional state such as anger, happiness, sadness, etc. In the view of Shivhare and Khethawat (2012), detecting a person's emotional state by analyzing a text document written by him or her may appear difficult, but it is often necessary because textual expressions are often not only direct expressions of emotion words, but also result from the interpretation of the meaning of concepts and the interaction of concepts described in the text document. A text-based emotion control system is imperative. It is a system that analyzes a text and detects emotions from the text by finding relations within the text and the emotions that lies within it using some emotional keywords, from nouns, verbs and adjectives. Thus, if a negative sentiment is presented in a text, the system detects it, and relays the result to the user. The system provides this text-based emotion control from a large number of the English lexicons. It is intended to serve as a viable way to protect humans from reading unprepared negative text-based messages which may hold certain dangerous health reactions. Over the years, the health issues have been on the high side due to the effect of reading text messages with negative emotions in them. To show the prevalence of the health issues caused by the reading text messages, Shalaby et al. (2022) demonstrated Text2quit and Quit4baby, two texting service programs to help people cease from smoking addiction; and Text4Mood and Text4Support, two other mobile text messaging programs to help people with mood disorders and other mental health conditions, including alcohol use disorder. Health issues such as heart attack, high blood pressure, hypertension, stroke and even bowel disorder can occur due to negative messages as noted by Ogbuju et al (2021) which also demonstrated that having a system that filters and block negative emotions in text messages from getting to the user will be of great help in reducing the impact of these health matters on our communities.

While humans are generally more adept at identifying emotions in text documents or blogs, this process can be time-consuming as it requires a thorough analysis of the text to determine the emotions being conveyed (Adiyanto, et al 2021). Therefore, there is a need for machine learning techniques to automate this task. Various machine learning approaches have been utilized over time to detect emotions in textual data. Shivhare and Khethawat (2012) conducted a knowledge-based survey on emotion detection using textual data and introduced a new architecture for recognizing emotions in text documents. This proposed architecture consists of two primary components: an emotion ontology and an emotion detector algorithm. By taking a text document and the emotion ontology as inputs, the emotion detector system can identify one of six emotion classes, namely love, joy, anger, sadness, fear, and surprise, as its output.

Jain, Kumar, and Fernandes (2017) developed a hybrid framework to automatically detect emotions in multilingual text data. Their model utilized natural language processing techniques to extract emotions from texts and classify them according to Ekman's emotion models (anger, disgust, happiness, sadness, fear, and surprise) using SVM and NB. This approach outperformed other methods, achieving an accuracy of 72.1% in a multilingual scenario.

Byun and Lee (2010) introduced a system that can automatically generate gesture animations from SMS and stylize them according to the emotion being conveyed. One of the key features of this system is its real-time algorithm, which enables the combination of gestures with emotions while guaranteeing a performance of 15 or more frames per second. Initially, the system extracts words that express feelings and their corresponding gestures statistically, and then employs Laban Movement Analysis theory to combine the gestures with the emotions.

Another related work produced the EmoHeart, a lexical rule-based system developed by Neviarouskaya et al. (2010) which detects emotions from text and visualizes them in a virtual environment. The system first searches for emotional abbreviations and emoticons; if none are found, it processes the sentence on various levels (such as word, phrase, and sentence levels) to create an emotional vector of the sentence, where each element represents the intensity of an emotional class. They achieved an average accuracy of 75% when testing the system on a manually annotated dataset. However, their method has some limitations, such as the inability to handle negation in sentences and the reliance on an affective database where emotions and their intensities were assigned manually to each word, making it challenging to extend their approach to classify additional emotions.

Danisman and Alpkocak (2008) investigated the automatic classification of emotions in text using the International Survey on Emotion Antecedents and Reactions (ISEAR dataset). They utilized the Vector Space Model (VSM) to classify 801 news headlines from the "Affective Task" in the SemEval 2007 workshop, which focuses on the classification of emotions and valences in text. In their study, they compared the findings of VSM classification with ConceptNet and advanced text-based classifiers like Naive Bayes and Support Vector Machines (SVM). The results indicated that VSM outperformed the other classifiers in detecting emotions in texts.

According to Ho and Cao (2012), emotions are connected to human mental states triggered by emotional events. They proposed a Hidden Markov Model (HMM) that applies this concept by considering each sentence as comprising multiple sub-ideas, where each idea represents an event that leads to a particular mental state. The system analyzes the sequence of events within a sentence to determine the most probable emotion conveyed by the text. However, the system's performance was limited due to the lack of consideration of the sentence's semantic and syntactic structures, making it non-context-sensitive. When tested on the ISEAR dataset, the system achieved an F-score of 35%, with the best precision being 47%.

Anusha and Sandhya (2015) proposed a hybrid approach that combined rule-based and learning-based methods. Their approach involved defining rules based on the syntactic and semantic features of text extracted through Support Vector Machine (SVM). Natural Language Processing techniques were employed to enhance the performance of learning-based classifiers, such as Naïve Bayes (Chen et al., 2020). The Ekman emotional model and ISEAR database were utilized in their study.

The reviewed literature reveals that different techniques have been employed for detecting emotions from text using machine learning. However, a common trend is the use of the ISEAR dataset in many of the models and projects developed for predicting emotions in text. Additionally, the literature has mainly concentrated on detecting emotions from blogs, websites, and news headlines.

The aim of this study is to develop a text-based emotion control model for short messaging services (SMS). To achieve this aim, the study designed an emotion framework that detects positive and negative messages, built a machine learning model that detects positive and negative messages and then returns feedback to the user or receiver of the text and evaluated the performance of the model. The identification of human emotions holds great importance in fields such as security and healthcare. It allows for the detection of human sentiments without the need for direct inquiry. This study therefore proposes a model for controlling the emotions of users by providing them with the emotional content of SMS messages before they are read. This would enable the blocking of text messages that contain negative emotions, preventing health issues such as high blood pressure and bowel disorders, which can be caused by abruptly reading emotionally charged messages. While there has been extensive research in the area of emotion detection using speech and facial recognition, little attention has been given to text-based emotion detection, specifically through SMS. This research aims to fill that gap by developing a machine learning model that detects emotions in SMS messages and relays the results to the message's recipient. The technical aspects of the project involve techniques such as sentiment analysis and opinion mining using lexicon-based approaches and machine learning algorithms to model the sentiments of texts. These techniques have been applied in previous works by Ogbuju et al (2022) to classify social protest messages in Emeka et al (2023) to review approaches for the control of fake news. The contribution of this study its potential to improve the well-being of individuals by providing them with the ability to control the emotional impact of textual messages.

2. Method

The system is designed to mitigate the negative effects of text messages on users by using a text-based emotion control system. Typically, SMS messages are delivered directly to the recipient without being filtered for positive or negative content. The message is transmitted from the sending device to the nearest cell tower, which forwards it to a Short Message Service Center (SMSC). The SMSC then relays the message to a cell tower near the receiving device, which then passes the message to the recipient. However, with the text emotion system in place, the message is intercepted at the receiving device, and if positive emotion is detected, it is delivered to the user. If negative emotion is detected, the message is blocked, and a notification is sent to the user to inform them that a message with negative emotion was received but has been blocked.

The main components of the system include an emotion detector and classification models which processes an input message into different emotion classes as output with which the message access is determined. The process framework is shown in Fig 1.

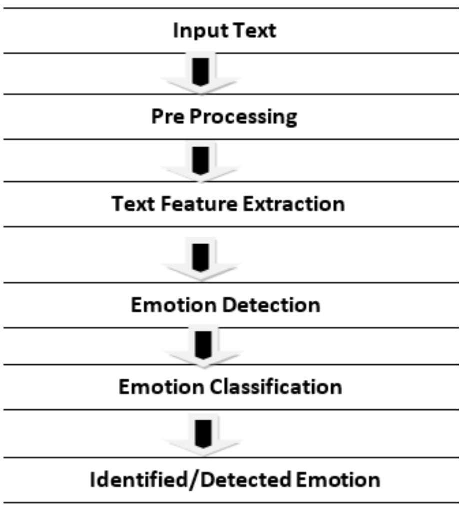
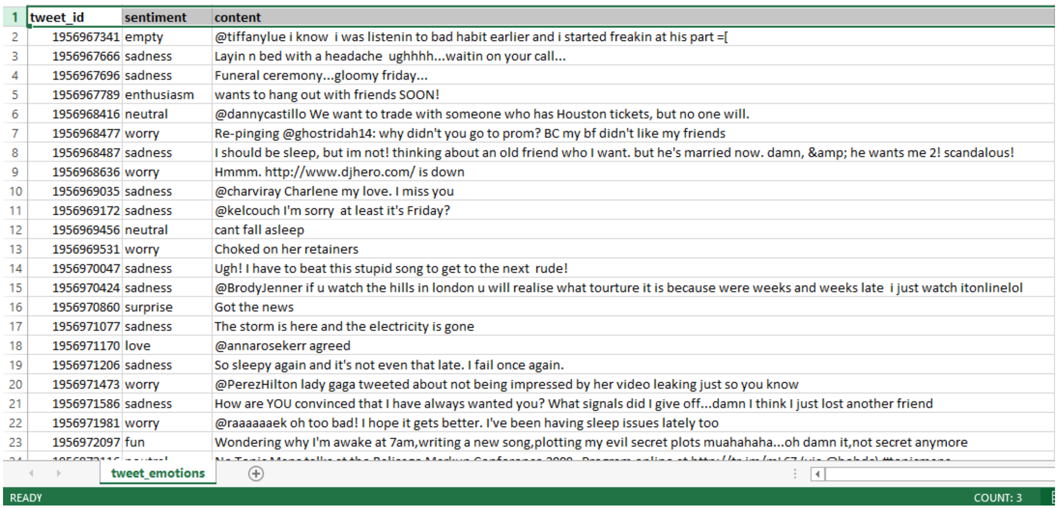


Fig 1. Process framework

The system consists of two major parts, the application interface development using Tkinter Python (Yadav & Raheman, 2023) and the detection/classification model development using Python libraries (Herath et al., 2020) such as Pandas (Marins et al., 2021), Numpy (Khamparia et al., 2020), Sci-kit Learn (Christ et al., 2018), Matplotlib (Imtyaz et al., 2020), and Seaborn (Nasiboglu & Nasibov, 2022). As shown in the process flow in Figure 1, six (6) functional modules, namely text input, text-preprocessing, feature extraction, emotion detection, classification and result display are followed in the model development. The Agile software development process was adopted for the application system, allowing for flexible changes and improvements to the application during the development process.

The system's user interface design consists of three distinct pages: the homepage, message received page, and message blocked page (Adebowale et al., 2019). The homepage enables the user to input text and send it via a button (Xu et al., 2015). The message received page displays the message to the recipient for reading (Oeldorf-Hirsch & Sundar, 2015). In contrast, the message blocked page shows a notification that the text has been blocked if the provided text contains negative emotions, preventing the recipient from reading it (Wu et al., 2020). The screenshot of the experimental dataset used in building the model is shown in Fig 2. It is a collection of tweets which had been labeled with the different emotions they convey.



	tweet_id	sentiment	content
2	1956967341	empty	@tiffanylue i know i was listenin to bad habit earlier and i started freakin at his part =[
3	1956967666	sadness	Layin n bed with a headache ughhhh...waitin on your call...
4	1956967696	sadness	Funeral ceremony...gloomy friday...
5	1956967789	enthusiasm	wants to hang out with friends SOON!
6	1956968416	neutral	@dannycastillo We want to trade with someone who has Houston tickets, but no one will.
7	1956968477	worry	Re-pinging @ghostridah14: why didn't you go to prom? BC my bf didn't like my friends
8	1956968487	sadness	I should be sleep, but im not! thinking about an old friend who I want. but he's married now. damn, & he wants me 2! scandalous!
9	1956968636	worry	Hmmm. http://www.djhero.com/ is down
10	1956969035	sadness	@charviray Charlene my love. I miss you
11	1956969172	sadness	@kelcouch I'm sorry at least it's Friday?
12	1956969456	neutral	cant fall asleep
13	1956969531	worry	Choked on her retainers
14	1956970047	sadness	Ugh! I have to beat this stupid song to get to the next rude!
15	1956970424	sadness	@BrodyJenner if u watch the hills in london u will realise what tourture it is because were weeks and weeks late i just watch itonlinelol
16	1956970860	surprise	Got the news
17	1956971077	sadness	The storm is here and the electricity is gone
18	1956971170	love	@annarosekerr agreed
19	1956971206	sadness	So sleepy again and it's not even that late. I fail once again.
20	1956971473	worry	@PerezHilton lady gaga tweeted about not being impressed by her video leaking just so you know
21	1956971586	sadness	How are YOU convinced that I have always wanted you? What signals did I give off...damn I think I just lost another friend
22	1956971981	worry	@raaaaaaek oh too bad! I hope it gets better. I've been having sleep issues lately too
23	1956972097	fun	Wondering why I'm awake at 7am,writing a new song,plotting my evil secret plots muahahaha...oh damn it,not secret anymore

Fig 2. Dataset

3. Results and Discussion

The Logistic Regression algorithm gave the highest accuracy of 98.4% among others as shown in Table 1. This was used to build the detection system.

Table 1. Model Evaluation

Model	Accuracy (%)	F1 Score (%)
SVM	91.7%	93%
NV	100%	73%
LR	98.4%	93%

The system's Graphical User Interface was designed using HTML and CSS. It provides the environment that makes the user's interaction with the system simple and efficient for accomplishing user goals (Cernea & Kerren, 2015; Stephens-Fripp, 2016). It also shows how the user performs certain interactions and the usability of the interface design. It serves as the screen of the sender of a text message (Fig 3).

Fig 3. Text Message Interface

When a user inputs a message with a positive emotion and clicks on the send button, the message is analysed analysed by the Text Based Emotion Detection System to give an output. The output displayed would be "Message Received" and the sent message would be allowed to be displayed. This is because this message will have no negative result or effect on the reader of the given message. This is shown in Fig 4.

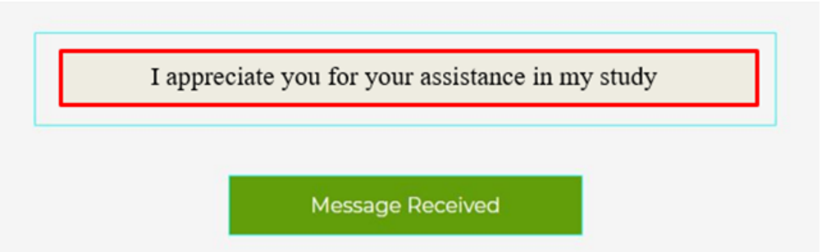


Fig 4. Positive message reception

When a negative message "I hate you and everyone around you" is inputed, it returns a blocked result as shown in Fig 5.

Fig 5. Negative message blockage

While previous works reviewed in this study concentrate on detecting emotions, this work had extended this feature to an application that would block negative emotions from getting to recipient. This work significantly impacts the field of health and well-being by addressing the health risks associated with negative emotions in text messages. It presents an innovative solution through machine learning and real-time emotion detection in SMS, with potential applications in improving users' emotional well-being and overall health. However, the study is subject to some limitations because the datasets used are labeled by individuals, and people may have varying interpretations and readings of emotions. Also, sarcasm in text can make it difficult to interpret emotions accurately since it may not be expressed directly. Again, word ambiguity can also pose a challenge in assigning keywords to a particular emotion class since a word can have different meanings in different contexts. In the current system, the model may be unable to recognize emotions in the absence of emotion keywords, which can lead to underestimation of emotions conveyed in a text.

4. Conclusion

The detection of emotions from textual data is crucial in addressing health concerns. To achieve this, a Text Based Emotion System was developed, which analyzes specific keywords in the text to determine its emotional content. The system developed in this project comprises two main components: a user interface designed using Python and a machine learning model that serves as the backend. The interface enables seamless interaction between the user and the model, while the machine learning model achieves an accuracy of 98.4% using Logistic Regression, making it the best option for emotion detection in textual data. This system can greatly reduce negative health outcomes resulting from exposure to negative emotions. Individuals with conditions such as bowel disorders and heart issues, which are exacerbated by negative messages, can benefit from this detection system. Future research should focus on improving the system to provide positive feedback to users when negative emotions are detected, as well as extending its application to social media and other contexts.

Author Contribution: All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding: This research received no external funding

Conflicts of Interest: Declare conflicts of interest or state, “The authors declare no conflict of interest.”

References

Adebowale, M. A., Lwin, K. T., Sánchez, E., & Hossain, M. A. (2019). Intelligent web-phishing detection and protection scheme using integrated features of Images, frames and text. *Expert Systems with Applications*, *115*, 300–313. https://doi.org/10.1016/j.eswa.2018.07.067

Adiyanto, O., Mohamad, E., Jaafar, R., & Faishal, M. (2023). Identification of Musculoskeletal Disorder among Eco-Brick Workers in Indonesia. International Journal of Occupational Safety and Health, 13(1), 29-40. https://doi.org/10.3126/ijosh.v13i1.44575

Anusha, V., & Sandhya, B. (2015). A learning-based emotion classifier with semantic text processing. In: El-Alfy, ES., Thampi, S., Takagi, H., Piramuthu, S., Hanne, T. (eds) Advances in Intelligent Informatics. Advances in Intelligent Systems and Computing, 320. Springer, Cham. https://doi.org/10.1007/978-3-319-11218-3_34

Byun, H. W., & Lee, J. S. (2010). Emotion-based gesture stylization for animated SMS. Journal of Korea Multimedia Society, 13(5), 802-816.

Cernea, D., & Kerren, A. (2015). A survey of technologies on the rise for emotion-enhanced interaction. *Journal of Visual Languages & Computing*, *31*, 70–86. https://doi.org/10.1016/j.jvlc.2015.10.001

Chen, S., Webb, G. I., Liu, L., & Ma, X. (2020). A novel selective naïve Bayes algorithm. *Knowledge-Based Systems*, *192*, 105361. https://doi.org/10.1016/j.knosys.2019.105361

Christ, M., Braun, N., Neuffer, J., & Kempa-Liehr, A. W. (2018). Time Series FeatuRe Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). *Neurocomputing*, *307*, 72–77. https://doi.org/10.1016/j.neucom.2018.03.067

Danisman, T., & Alpkocak, A. (2008, April). Feeler: Emotion classification of text using vector space model. In AISB 2008 convention communication, interaction and social intelligence, 1(4): 53-59

Emeka, O., Taiwo, A., Francisca, O. (2023). Text analytics solutions for the control of fake news: Materials and methods. International Journal of Open Information Technologies, 11(3): 69-74

Herath, U., Tavadze, P., He, X., Bousquet, E., Singh, S., Muñoz, F., & Romero, A. H. (2020). PyProcar: A Python library for electronic structure pre/post-processing. *Computer Physics Communications*, *251*, 107080. https://doi.org/10.1016/j.cpc.2019.107080

Ho, D.T., & Cao, T.H. (2012). A high-order hidden Markov model for emotion detection from textual data. In: Richards, D., Kang, B.H. (eds) Knowledge Management and Acquisition for Intelligent Systems. PKAW 2012. Lecture Notes in Computer Science, 7457. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-32541-0_8

Imtyaz, A., Abid Haleem, & Javaid, M. (2020). Analysing governmental response to the COVID-19 pandemic. *Journal of Oral Biology and Craniofacial Research*, *10*(4), 504–513. https://doi.org/10.1016/j.jobcr.2020.08.005

Jain, V. K., Kumar, S., & Fernandes, S. L. (2017). Extraction of emotions from multilingual text using intelligent text processing and computational linguistics. Journal of computational science, 21, 316-326.

Khamparia, A., Singh, A., Luhach, A. Kr., Pandey, B., & Pandey, D. K. (2020). Classification and Identification of Primitive Kharif Crops using Supervised Deep Convolutional Networks. *Sustainable Computing: Informatics and Systems*, *28*, 100340. https://doi.org/10.1016/j.suscom.2019.07.003

Krishnan, H., Elayidom, M. S., & Santhanakrishnan, T. (2017). Emotion detection of tweets using Naïve Bayes classifier. International Journal of Engineering Technology Science and Research, 4(11), 457-462.

Marins, M. A., Barros, B. D., Santos, I. H., Barrionuevo, D. C., Vargas, R. E. V., de M. Prego, T., de Lima, A. A., de Campos, M. L. R., da Silva, E. A. B., & Netto, S. L. (2021). Fault detection and classification in oil wells and production/service lines using random forest. *Journal of Petroleum Science and Engineering*, *197*, 107879. https://doi.org/10.1016/j.petrol.2020.107879

Nasiboglu, R., & Nasibov, E. (2022). FyzzyGBR—A gradient boosting regression software with fuzzy target values. *Software Impacts*, *14*, 100430. https://doi.org/10.1016/j.simpa.2022.100430

Neviarouskaya, A., Prendinger, H., & Ishizuka, M. (2010). EmoHeart: conveying emotions in second life based on affect sensing from text. Advances in Human-Computer Interaction, 1.

Ogbuju, E., Yemi-Peters, V., Osumah, T., Agbogun, J., Ejiofor, V. (2021). A frontpage online news analysis: An opinion mining approach. Journal of Computer Science and Its Application, 28(1): 107-120. https://doi.org/10.4314/jcsia.v28i1.9

Ogbuju, E., Mpama, I., Oluwafemi, T.M., Ochepa, F.O., Agbogun, J... (2022). The sentiment analysis of EndSARS protest in Nigeria. Journal of Applied Artificial Intelligence, 3(2): 13-23

Oeldorf-Hirsch, A., & Sundar, S. S. (2015). Posting, commenting, and tagging: Effects of sharing news stories on Facebook. *Computers in Human Behavior*, *44*, 240–249. https://doi.org/10.1016/j.chb.2014.11.024

Shalaby, R., Adu, M. K., El Gindi, H. M., & Agyapong, V. I. O. (2022). Text Messages in the field of mental health: Rapid review of the reviews. Frontiers in Psychiatry, 13, 921982. https://doi.org/10.3389/fpsyt.2022.921982

Shivhare, S. N., & Khethawat, S. (2012). Emotion detection from text. https://doi.org/10.48550/arXiv.1205.4944

Stephens-Fripp, B. (2016). Combining Local and Global Features in Automatic Affect Recognition from Body Posture and Gait. *University of Wollongong Thesis Collection 1954-2016*. https://ro.uow.edu.au/theses/4923

Wu, D., Moody, G. D., Zhang, J., & Lowry, P. B. (2020). Effects of the design of mobile security notifications and mobile app usability on users' security perceptions and continued use intention. *Information & Management*, *57*(5), 103235. https://doi.org/10.1016/j.im.2019.103235

Xu, J., Kang, Q., Song, Z., & Clarke, C. P. (2015). Applications of Mobile Social Media: WeChat Among Academic Libraries in China. *The Journal of Academic Librarianship*, *41*(1), 21–30. https://doi.org/10.1016/j.acalib.2014.10.012

Yadav, R., & Raheman, H. (2023). Development of an artificial neural network model with graphical user interface for predicting contact area of bias-ply tractor tyres on firm surface. *Journal of Terramechanics*, *107*, 1–11. https://doi.org/10.1016/j.jterra.2023.01.004