Al-Rafidain Journal of Computer Sciences and Mathematics (RJCSM)

www.csmj.uomosul.edu.iq

Harnessing Machine Learning to Forecast Social Media Advocacy Success

Bala Muhammad Muhammad[1], Abdulmalik Ahmad Lawan[2] and Abdo Sulaiman Abdi[3]

[1] Department of Mathematics and Data Science, Andalas University, Padang, Indonesia
[2] Department of Computer Science, Aliko Dangote University of Science and Technology, 713281 Wudil, Nigeria
[3]Information Technology Department, Duhok Technical Institute, Duhok Polytechnic University, Duhok, Iraq
Email: muhammadbala466@gmail.com[1] , aalawan@kustwudil.edu.ng[2] , and abdo.abdi@dpu.edu.krd[3]

Article information

Article history:
Received 26 March, 2026
Revised 30 April, 2026
Accepted 6 May, 2026
Published 25 June, 2026

Keywords:
#JusticeForNavalAbbas,
Advocacy,
Activism,
Social Media,
Machine Learning.

Correspondence:
Abdo Sulaiman Abdi
Email:
abdo.abdi@dpu.edu.krd

Abstract

Social media has modernized advocacy by enabling individuals to mobilize, spread opinions, and champion justice globally. Accurate prediction of social media advocacy campaign success can play a significant role in promoting social and political change. Motivated by the recent #JusticeForNavalAbbas advocacy campaign in Nigeria, this study aims to predict the success rate of posts related to the campaign. By examining how social media platforms including Facebook, TikTok, Instagram, and X (formerly known as Twitter) influence public opinion, this paper investigates how these platforms attract engagement, amplify the voices of the marginalized, and encourage collective action. The present research leverages five machine learning techniques, including Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and XGBoost (XGB). The dataset was preprocessed and split into training and testing subsets comprising 75% and 25% of the data, respectively. Model performance was evaluated using Accuracy, Precision, Recall, F1-Score, and Area Under the ROC Curve (AUC). RF and XGB demonstrated exceptional performance in predicting social media post success, achieving accuracies of 98% and 97%, respectively. The findings highlight the significance of social media platforms in amplifying advocacy campaigns at both local and global levels.

DOI: 10.33899/rjcsm.v20i1.60662,

1. Introduction

Social media platforms are tools that enable people to interact freely with others and provide various ways for marketers to reach and engage with consumers [1]. Nowadays Social media platforms have upgraded above their initial use as platforms that enable individuals to communicate and interact online by turning into a strong weapon for promoting social and political change with real-time access to information. Social media provide room for the masses to question authorities and demand justice. It does so by promoting the opinions of the underprivileged in demand for justice. The power of social media for advocacy has been proved by notable movements like Nigeria's #JusticeForMohbad [2], #EndSARS, which campaigned against police brutality [1], [2], [3], and the international #EndBadGovernments campaigns across African countries and other continents. The study of Ogunleye and Agber [4] examine how X (formerly Twitter) played a pivotal role in mobilizing the #EndSARS protest, amplifying collective outrage, and facilitating communication across diverse communities. Their study highlights X's critical function in shaping contemporary digital activism and the networked public sphere.

A social media advocacy under the hashtag #JusticeForNavalAbbas for the case of Seaman Haruna, a naval officer who was allegedly facing long detention and mental illness problems, has become increasingly popular in Nigeria. Social media users have come together in favor of Seaman Abbas through this online campaign, appealing to the rightful authorities to give him justice and

accountability. The present instance serves as a prime instance showing how social media sites can expose structured inequalities while offering visibility to those who are unlikely to be heard. Comparable to earlier popular justice movements like #EndSARS and #JusticeForMohbad. The current #JusticeForNavalAbbas movement demonstrates how vital social media is for boosting public interactions, keeping authorities responsible, and strengthening equality for all. However, unlike the Western world, there is insufficient research on successful social media campaigns in Nigeria and Africa, at large, as shown in **Figure 1**.

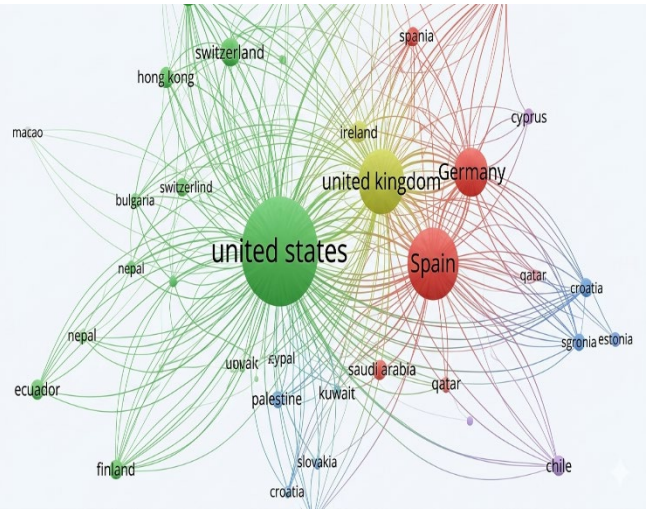


Figure 1. Country-based meta-analysis of studies on social media campaigns (Source: Scopus Database 2025).

In essence, despite the influence of social media in shaping public opinion and mobilizing social movements, there is a noticeable gab for predicting the success of online advocacy campaigns based on quantifiable factors because of unpredictable levels of engagements that leads to the success of a campaign while others remain unsuccessful. There is need for a modernized approach to efficiently analyze key engagement metrics such as likes, retweets, followers, and hashtags to predict the success of campaign posts.

Therefore, this work uses the #JusticeForNavalAbbas movement as a case study to evaluate the relationship between activism, technology, and social media. To better understand the processes that drive the success of social media activism, this study examines the use of machine learning to predict the success rate of the #JusticeForNavalAbbas campaign by analyzing key factors that attract post engagements and strategies for digital activism campaigns to achieve higher engagements. The following sections of this article explore related works, tools and materials, and conclusions, along with recommendations for future improvements.

2. Review of Related Works

2.1 The Transformative Role of Social Media in Empowerment, Education, Advocacy, and Engagement

Existing literature has explored the impact of social media platforms as powerful advocacy tools across various sectors, including non-profit organizations, community health, and political activism, among others, yielding remarkable achievements [5], [6], [7], [8], [9], [10], [11], [12]. In the work of D. Farsi [6], an extensive literature review of 158 research articles was carried out to explore how social media functions as a vital tool in healthcare for purposes such as health promotion, professional networking, education, and public health engagement. The review concluded that combining social media with traditional healthcare communication methods enhances communication, supports professional development, and ultimately improves service effectiveness. Moreover, Ahmad S.A. and Joseph A.B. [5], study looked at how social media influences and energizes women for community development and found that social media greatly increases women's involvement, consciousness, and partnership in the creation of community development. Similarly, Kumari [7], utilized qualitative methods to investigate the role of social media in empowering women's social, psychological, and financial aspects, with a focus on rural and urban women entrepreneurs in India. One of the major findings of the study was that social media is a powerful agent for gender equality. This is made possible by the loudspeaker effect, which gives women's voices more reach, by the spread of digital literacy among women, and by the significant influence of social media on the regulatory process. Furthermore, Makatlal V. and Kumar D. researched the impact of social media on political mobilization and activism, stating that it helps to bring out the voices of the marginalized groups, supports the formation of groups for common goals, and changes the way people get involved in socio-politics nowadays. However, the authors also acknowledge that there are some problems such as fake news and users staying in their own ideological bubbles [8]. In addition, the investigation performed in the work of Michelle et al. [9], content analysis and virtual ethnography was employed to understand the adaptation of non-profit organizations to new media by their change in online communication from a one-way information sharing to a two-way interactive communication. The research revealed that social media facilitates participatory advocacy, enhances the bond with supporters, and inspires collective action both online and offline. Also, the study of Michelle at al. [10], focused on the Bersih movement in Malaysia in order to demonstrate how social media enables citizens to organize collective action and influence the change of institutions. One of its findings was that the use of social media is not simply limited to spreading information, but it actually has a significant influence in the development of social

movements and in finding solutions for complicated civic and political issues. Besides, Silvia et al. [11], employed a causal design and structural equation modeling on data gathered from 211 small and medium-sized enterprises in the Brazilian healthcare sector to identify the factors that influence the use of social media. The study revealed that organizations are motivated to make such a move by the advantages they perceive social media to bring and the characteristics of the organization, whereas the surroundings or the environment had only a negligible effect. As well as, the research conducted by Jennifer et al. [12], looked into the content of tweets from the 2014 U.S. gubernatorial campaigns to understand how candidates used social media to attack their rivals. One of the findings was that the challengers were more aggressive in their attacks on social media, however, the attacks of the incumbents were more popular as they got shared more often among the users of social media. Likewise, Ismi et al. [13], employed a descriptive quantitative approach with 100 respondents to assess the impact of the Spotify Wrapped digital campaign on brand advocacy across Instagram, X, and TikTok. Results showed a strong positive correlation ($r = 0.822$, $p < 0.01$), indicating that higher user participation in the campaign significantly enhances Spotify's brand advocacy and active user promotion on social media.

Interestingly, social media has had a significant impact on education and marketing [1], [14], [15]. In the work of Appel et al. [1], the research delved into the changing function of social media in marketing by figuring out nine new themes from the viewpoints of consumers, the industry, and public policy. The paper, based on academic insights and expert discussions, traced the impact of these trends on the future of social media interaction, planning, and investigation in the digital market. Moreover, Cavus et al. [14], did a systematic review of the use of Social Networking Sites (SNSs) in education during the COVID-19 pandemic, highlighting their function in maintaining learning when there were difficulties in eLearning. The research figured out that SNSs facilitate communication, engagement, and education continuity, thus, they are a source of feasible methods for schools, teachers, and students to prevail over the crisis caused by the pandemic. Nonetheless, Van Den Beemt et al. [15], has gone through 271 articles to figure out the differences in the way the social media integration in the education through three-level curriculum model had been presented in the studies, and the outcomes of these studies. The main point of the paper is that such use of the media depends on the culture of the school, the attitude of the teachers, the support, and the professional development of teachers, and underlines the importance of the organization of the integration and having clear learning objectives for the educational impact to be increased.

Notably, social media support movements like #BlackLivesMatter in the U.S. and #EndSARS in Nigeria have gained influence globally [16], [17]. Particularly, Ghozian A.P. and Syaifa T. [17], employed content analysis of Twitter and Instagram posts to investigate the presence of the concept of hyperreality in the #BlackLivesMatter movement. It showed that although social media facilitated worldwide activism, a large part of the interaction became so far from the movement's fundamental message. The study ended up asserting that one of the effects of hashtag activism is that real social issues get turned into the so-called 'performative digital expressions' which are more about gaining visibility than doing justice. Similarly, Dambo et al. [16], conducted research on the #EndSARS movement in Nigeria as a case to determine how social media enables youth activism against police brutality and at the same time points out its limitations in a repressive political environment. The study found that social media, in spite of being a platform that loudens the voices of the oppressed and incites collective action, the existence of structural and political barriers which obstruct the accomplishments of such online movements. In essence, there are numerous domains on social media campaigns, to explore the popular domains, a meta-analysis of the most common terminologies from relevant literature is depicted with the help of **Figure 2**.

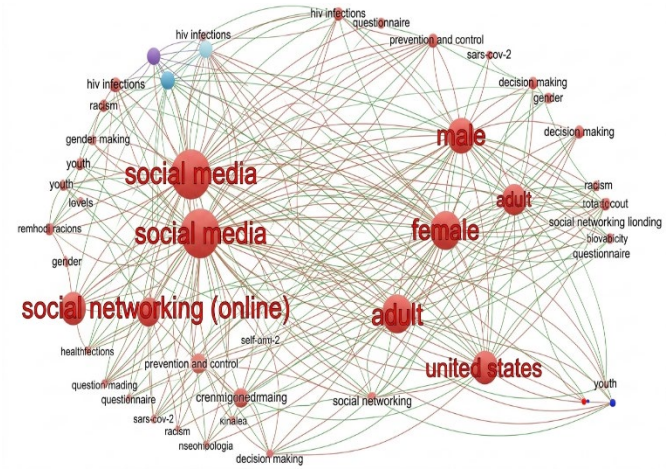


Figure 2. The key terms on social media campaigns extracted from 4,661 records between 2008 and 2025 (Source: Scopus 2025).

To recap, the studies examined reveal that social media, as a major factor, has a multifaceted impact on the aforementioned areas: empowerment, education, activism, healthcare, and marketing. The use of social media ranges from community mobilization and giving loudspeakers to those in the margins to improving learning and brand engagement. It is at once the cause of the change and the means of linking up. Nevertheless, its capabilities are frequently questioned by problems like fake news, performative activism, and digital divide. In essence, social media still constitutes a potent source of engagement, knowledge, and creativity in different fields of society.

2.2 Machine Learning Applications in Education, Healthcare, Careers, and Marketing

Machine learning techniques have been used in various sectors, highlighting their diverse applications in healthcare, education, and business amongst others [18], [19], [20], [21], [22], [23], [24], [25]. In the work of Buenaño-Fernández et al. [18], machine learning techniques were applied to predict students' final grades based on historical academic records from a computer engineering program at a university in Ecuador. Through data preprocessing, clustering, and supervised learning, the findings revealed that machine learning is a powerful tool for predicting academic performance. Moreover, AA Lawan et al. [19], utilized machine learning application in the detection of autism spectrum disorder (ASD) and recognizes significant results of the use of imaging and behavior data. The paper discusses the issue of the distance between data-driven models and medical practice that needs to be overcome in order to have dependable diagnostic tools that can be used in the real world. Furthermore, Cavus at al. [20], The study examined various applications of machine learning in the diagnosis of autism spectrum disorder (ASD) and concluded that the most promising results were those achieved by using imaging and behavioral data. In addition, Hebehh et al. [21], explored various machine learning applications in healthcare and discussed the most important algorithms, main usages, as well as ethical and privacy issues that were put forward. Also, Jhaveri et al. [22], research looked into the past Kickstarter projects to judge whether they would be successful or not, thus providing the owners of these projects the means to plan efficiently. By means of classification and boosting algorithms, the study concluded that the best prediction accuracy was achieved by the combination of Weighted Random Forest with AdaBoost on a subsampled dataset. Additionally, Lee L. and Shin Y. [23], proposed a unified framework for the residual learning and random forest regression for social media prediction tasks with the primary goal of predicting the view count of the post. Besides, Rahmayanti et al. [24], utilizes different machine learning classifiers to predict the fetus health including, Artificial Neural Network (ANN),

Long-short Term Memory (LSTM), XG Boost (XGB), Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Light GBM (LGBM), and Random Forest (RF). As well as, Bala et al. [25], applied machine learning models to student's dataset by merging academic performance and personal attributes, to predict the most suitable careers for Nigerian students. XGBoost reached the highest accuracy (85.7%), indicating that the use of personal and academic data together enhances the provision of career guidance even more for non-STEM directions and thus, it is a good source of information for students, counselors, and policymakers. In fact, Huang et al. [26], proposed the of use multi-aspect features combined with the random forest model for popularity predictions. Firstly, they extracted features by combining both the metadata of the posts and users' features. Indeed, [27], conducted a sentimental analysis with data sources from Twitter using the Random Forest algorithm approach and recorded an accuracy of around 75%. Particularly, the work of Gabay et al. [28], utilizes various machine learning algorithms in predicting the success of marketing campaigns on social media. Moreover, machine learning techniques have enhanced healthcare and have recorded exceptional performance.

2.3 Machine Learning for Social Media Advocacy Forecasting

Worldwide multiple research works have delved into the utilization of machine learning methods for the prediction and assessment of the success of social media campaigns. These papers reveal a wide range of modeling strategies such as supervised learning, graph-based models, and deep learning architectures to anticipate engagement, virality, and advocacy results. For instance, the research of Zhixuan X. and Minghui Q. [29], Utilized a temporal-spatial cascade convolutional learning framework (ViralGCN) on Weibo retweet data, the study effectively integrated structural and temporal information to predict online content popularity, revealing that both broadcast and structural virality, along with early-stage cascade patterns and user timing, significantly influence viral success. Moreover, Macro et al. [30], Use interpretable machine learning approach on over 40,000 tweets from 23 major museums, the study developed a predictive model to assess tweet appreciation and suggest content improvements, revealing key tweet features that significantly enhance engagement and success on Twitter. Furthermore, Andariesta D. [31], on about 12,786 tweets from major Indonesian e-commerce platforms, the study applied multinomial logistic regression, decision tree, k-nearest neighbor, and random forest models to predict Twitter engagement, finding that interactivity, vividness, and timing strongly influence engagement, with the random forest model achieving the highest predictive performance. Additionally, Venkata [32], Analyzed social media advertisements to enhance audience classification and targeting accuracy, revealing that understanding behavioral factors enables marketers to optimize ad campaigns and significantly improve customer conversion rates. Interestingly, Biricik et al. [33], predicted the popularity of social media campaigns for an e-commerce company, demonstrating that AI-driven prediction effectively reduces unnecessary advertising, production, and material costs while significantly enhancing campaign efficiency. Particularly, in the work of Vashisht [34], extensive preprocessing, feature engineering, as well as various machine learning models was utilized to analyzed diversified Twitter data to develop a predictive model for predicting and proposing emerging trendy hashtags, which demonstrates that machine learning precisely identifies interaction patterns for predicting emerging trends.

Table 1. First few records of the dataset.

Users	Followers	Retweets	Likes	Hashtags
User_1	35900	54	1308	#JusticeForNavalAbbas
User_2	28500	23	41	#JusticeForNavalAbbas
User_3	7370	548	4083	#HumanRights
User_4	36545	454	4983	#JusticeForNavalAbbas
User_5	6496	724	1759	#SocialJustice

This study employs machine learning techniques, including Random Forest (RF), Logistic Regression (LR), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and XGBoost (XGB), to predict the success rate of social media posts related to the #JusticeForNavalAbbas movement. By identifying key success indicators, the research highlights how social media amplifies marginalized voices in the pursuit of justice, demonstrating its power as an advocacy tool.

3. Materials and Methods

The current study uses a machine learning approach, particularly the random forest to predict the success of social media posts under the hashtag #JusticeForNavalAbbas movement. The Dataset used was collected through a manual data extraction from assorted social media accounts, and pre-processed using Python libraries. The data collection was conducted from 15th to 19th September 2024 across Four social networking sites, viz: Facebook, TikTok, Instagram and X (formally known as Twitter). The dataset was split into training and testing subsets for validation. Performance was measured using accuracy, precision, recall, F1-score metrics. **Figure 3** explains the flowchart of the present study. In the present work, as highlighted in **Figure 3**, the methodology that guided the steps is described as follows:

a. Data Collection

The The dataset used in this research was collected through systematic observation of publicly accessible social media posts published during the #JusticeForNavalAbbas campaign. Data was manually extracted from four social networking platforms X (formerly known as Twitter), Facebook, TikTok, and Instagram between 15th and 19th September 2024, representing the peak activity period of the campaign. A purposive sampling approach was adopted, targeting only posts that explicitly contained the campaign hashtag or directly related hashtags, ensuring thematic consistency across the dataset. Posts were included if they contained the primary campaign hashtag, were publicly accessible, and contained complete engagement data. Posts were excluded if they belonged to private accounts, contained missing engagement data, or were identified as unrelated promotional content. Accounts displaying bot-like characteristics such as zero followers or repetitive posting patterns were manually reviewed and excluded. Following data cleaning, duplicate records and cases with missing values were removed using Python's Pandas library, resulting in a final dataset of 1,016 posts, each containing information on users' total Followers, Retweets, Likes, and used Hashtags, as shown in **Table 1**.

The #JusticeForNavalAbbas campaign was selected because it represents a recent, organic, and locally driven advocacy movement in Nigeria, a context significantly underrepresented in existing social media research, as highlighted in **Figure 1**.

The #JusticeForNavalAbbas campaign analysis revealed its main objectives through the examination of hashtag appearances within campaign-related social media content. Campaign participants show through their most used hashtags what main subjects and priorities they want to highlight. The hashtags used in the campaign is shown in **Figure 3**.

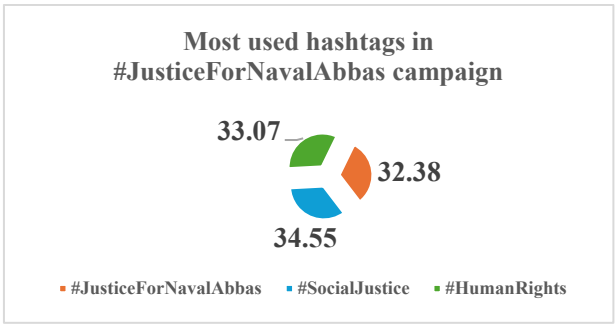


Figure 3. Most used hashtags in #JusticeForNavalAbbas campaign.

Figure 3. showcased the most used hashtags by social media users during the #JusticeForNavalAbbas campaign. Moreover, **Figure 4**. illustrates the proportion of Likes to Retweets for posts during the #JusticeForNavalAbbas campaign.

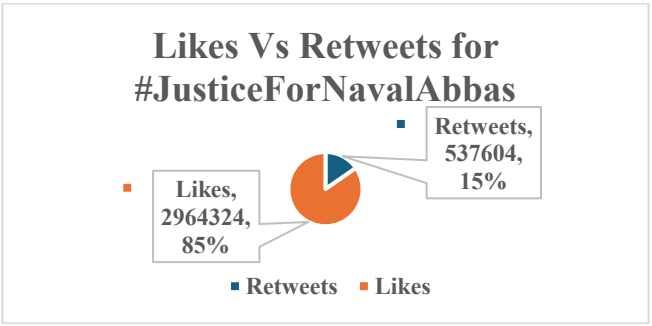


Figure 4. Likes VS retweets distribution in the #JusticeForNavalAbbas Campaign.

Figure 4. revealed that Likes take a larger portion with (85%) and 15% of retweets during the #JusticeForNavalAbbas campaign as such making it a strong predictive feature in our proposed model.

b. Data Preprocessing

After the data acquisition stage, the data preprocessing stage commenced with data cleaning, which involved removing duplicated records and cases with missing crucial inputs to

ensure accurate and reliable prediction results. Python libraries, specifically Pandas and NumPy, alongside Microsoft Excel, were utilized as the primary tools for data processing in the current study.

A binary target variable, 'Success,' was constructed to classify each post as either successful (1) or unsuccessful (0). To mitigate the risk of target leakage, it is important to clarify the distinction between the features and the target variable. While Likes and Retweets appear in both the feature set and the success criterion, they serve fundamentally different roles. As features, they represent the observed engagement behavior of individual posts during the campaign window. As part of the success label, they function as a threshold-based classification instrument used to categorize posts into high-impact and low-impact classes, a post-hoc labeling practice that is consistent with engagement-based classification approaches widely adopted in social media analytics research. Specifically, a post was classified as successful (1) if it recorded Likes ≥ 1000 or Retweets ≥ 500, reflecting a level of engagement that meaningfully exceeded the general interaction observed across the dataset, and unsuccessful (0) otherwise. This design is consistent with studies that use engagement metrics both as predictors and as outcome proxies when temporally concurrent data represents the only available source, particularly in resource-constrained and manually collected datasets. Nevertheless, the authors acknowledge this as a methodological consideration, and recommend that future studies incorporate temporally separated or pre-campaign features — such as pre-campaign follower growth rates or historical posting frequency — to establish a stricter and more independent predictive boundary.

Furthermore, the Hashtags feature was label-encoded into a numerical column, Hashtags_encoded, using Scikit-Learn's LabelEncoder to make it compatible with the machine learning algorithms employed. The selected features for modeling were Followers, Retweets, Likes, and Hashtags_encoded. The entire dataset was then partitioned into two subsets: a training sub-dataset comprising 75% of the data and a testing sub-dataset comprising the remaining 25%, using Scikit-Learn's train_test_split function. The partitioned data was subsequently processed using five machine learning algorithms: Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and XGBoost (XGB). Model performance was evaluated using multiple evaluation metrics, including Precision, Recall, F1-Score, Area Under the ROC Curve (AUC), and Confusion Matrix, to ensure a comprehensive and balanced assessment of each algorithm's discriminative power and class-wise prediction quality.

c. Model Development

Machine learning is a powerful tool and widely used for its accurate result. Meanwhile, this study focuses on predicting the success of a social media campaign advocacy using machine learning methods. In the current work, Random Forest (RF) was used for it exceptional ability in handling feature interactions and consists of many decision trees [35]. Moreover, the Logistic Regression (LR) is widely known or its ability to interpret binary classification problems. Also, the Support Vector Machine (SVM) which is the most widely used state-of-the-art machine learning technique analyze data used for classification and regression analysis [36]. Furthermore, K-Nearest Neighbors (KNN) algorithm is a simple, supervised machine learning algorithm that can be used to solve both classification and regression problems [36]. Ultimately, XGBoost (XGB) is an efficient algorithm that processed data using gradient boosting. Moreover, validation of the models was done with a test-set proportion of 25% from overall dataset. All processes from training models, validation models to parameter optimization used Python's Scikit-Learn module. The model parameters used are presented in **Table 2**.

Table 2. Parameters used.

Algorithm	*Hyperparameters*
RF	Number of estimators, maximum depth, class weight, minimum samples split, and maximum features
LG	Regularization strength ((C)), penalty type, and solver
SVM	Kernel function, regularization parameter ((C)), gamma, and probability output
KNN	Number of neighbors, distance metric, and weight function
XGB	Number of estimators, learning rate, maximum depth, and subsample ratio

For each model, a set of hyperparameters were utilized for optimal performance, as shown in **Table 2**. They are employed to achieve, avoid overfitting and improve performance of the algorithms. For instance, 100 is used as estimation number for both Random Forest (RF) and XGBoost (XGB) classifiers with a max depth of 10, balanced class weight, min_samples_split was set to be 5 and max_features to be sqrt, thus preventing overfitting and promoting diversity within the model. For the case of Logistic Regression (LR) regularization is the most important feature with the parameter C was set to be 1.0, and liblinear assigned for solver, penalty was assigned a value of 12 to handle multicollinearity among features. In the Support Vector Machine (SVM) model, the value of `C` was also set at 1.0 to balances the trade-off between minimal training error and high margin. The kernel function employed was rbf (Radial Basis Function) which enable the model to learn the non-linear relationships among the features utilized. A parameter gamma was used and was set to scale, an automatic parameter that scales the effect of individual training samples by data variability. In addition, probability was set to True to enable probabilistic prediction, which would be needed to calculate evaluation metrics such as AUC-ROC. In the K-Nearest Neighbors (KNN) algorithm, the number of neighbors was set to 5, Euclidean as the distance metric and uniform as weight function. Lastly, for the XGBoost (XGB) classifier, gradient boosting was implemented using well-adjusted parameters with

estimators' number to be 100, learning_rate value as 0.1, and 6 as the max_depth, the subsample was limited at 0.8. thus, avoiding underfitting and overfitting. Finally, the chosen hyperparameters for all models were employed to enhance the prediction capacity of its respective algorithm.

d. Model Evaluation Metrics

The evaluation process tested Random Forest (RF) and Logistic Regression (LG) and Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) and XGBoost (XGB) classifiers by using a comprehensive assessment approach. The evaluation process used multiple metrics which included Accuracy and Precision and Recall and F1-Score and Area Under the Curve (AUC) to assess the performance of models in both discriminative power and class-wise prediction quality. **Figure 5** Present the summary of the methodology employed in the current study.

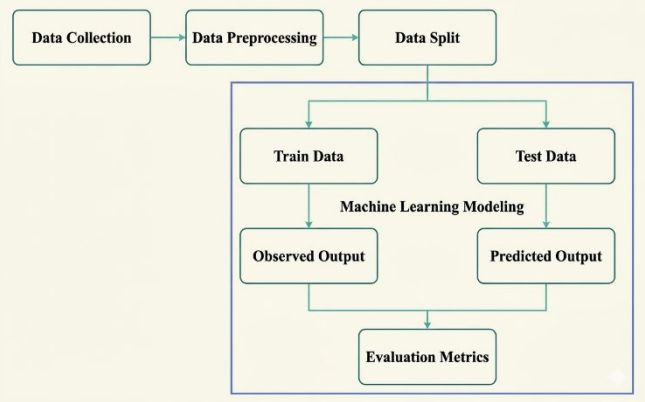


Figure 5. Workflow summary.

4. Results and Discussion

In the current study, Different machine learning algorithms were used to develop a prediction model in order to predict post success in social media, particularly for posts related to the #JusticeForNavalAbbas movement in Nigeria. The findings of the research demonstrated that the machine learning algorithm were fitted for the task. The present work was evaluated using different performance metrics, which include accuracy, precision, recall, and F1-score. Additionally, Correlation heatmap was utilized for insights into the relationships between the features used to predict the success of the #JusticeForNavalAbbas social media campaign, shown in **Figure 6**.

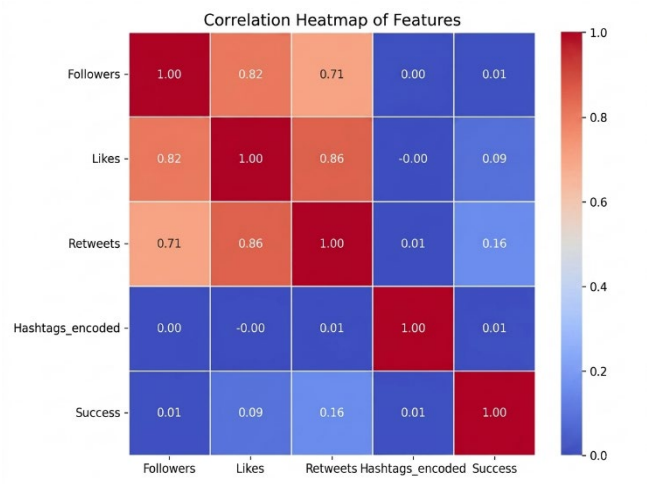


Figure 6. Correlation heatmap of features.

From the visualization **Figure 6**, it is evident that likes, retweets, and followers shown a strong positive correlation. This indicates that posts that receive more likes are also likely to be retweeted, which is expected in viral or successful campaigns. These features collectively reflect social media users' interest and interaction, thus making them strong predictors of campaign success. Though Followers contributed moderately, it alone was insufficient for strong prediction. Hashtags did not correlate linearly with success. Importantly, the heatmap ensures that the selected variables are informative, not redundant, and does not indicate any extreme multicollinearity in forecasting the success of social media campaign. Moreover, models like Logistic Regression (LR) learn from the data without being destabilized by overlapping information.

The classifiers used in the research namely, RF, LG, SVM, KNN, and XBG were evaluated using different evaluation metrics and models produce impressive model performance. Particularly the Random Forest (RF) has achieved an overall accuracy of 98% in predicting social media posts' success in the test dataset. Additionally, the precision, recall, F1-scores and AUC were 97%, 100%, 98%, and 99%, respectively. Followed by XGBoost (XGB) with an overall accuracy of 97%. Moreover, the Logistic Regression (LR), and Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) were the less performant models with overall accuracy of 89%, 87%, and 85% respectively, shown in **Table 3**.

Table 3. Result of classifiers.

Model	*Accuracy*	*Precision*	*Recall*	*F1-Score*	*AUC*
RF	0.98	0.97	1.00	0.98	0.99
LR	0.89	0.85	0.91	0.88	0.91
SVM	0.87	0.84	0.89	0.86	0.89
KNN	0.85	0.83	0.87	0.85	0.87
XGBoost	0.97	0.96	0.99	0.97	0.98

The study further demonstrates that the most important indicators for efficacy are posts with a high volume of likes and retweets, shown in **Figure 7**. The feature importance in predicting success rate also illustrates that likes, retweets, and followers are the key elements of the success prediction of a post. Moreover, users with more followers also attract higher engagement, which leads to the success of social media campaigns.

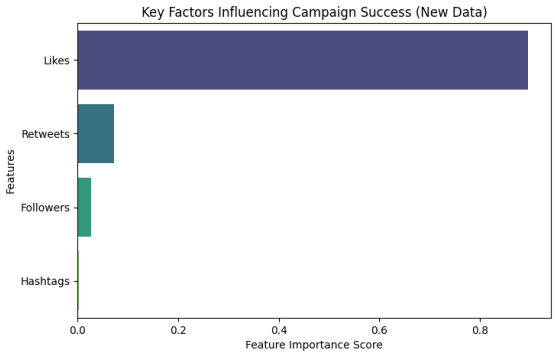
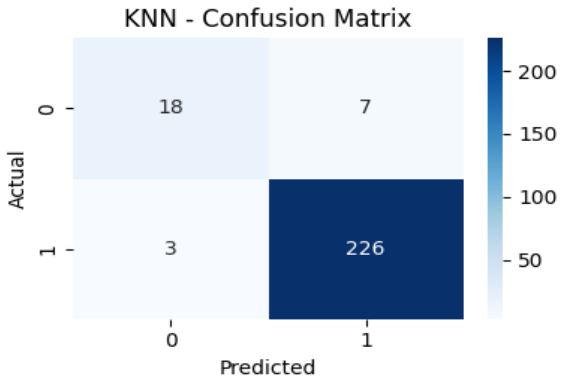
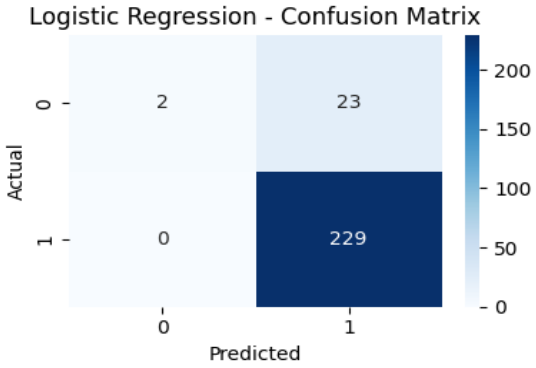


Figure 7. Feature importance plot.

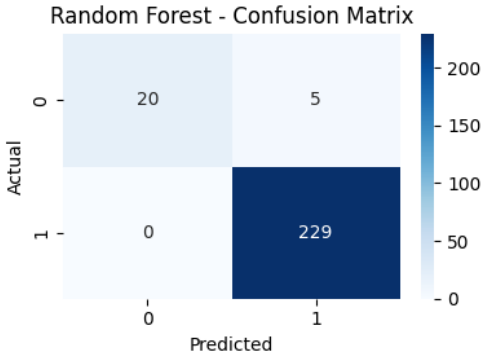
The confusion matrices below **Figure 7** provide a detailed breakdown of each model's classification accuracy.



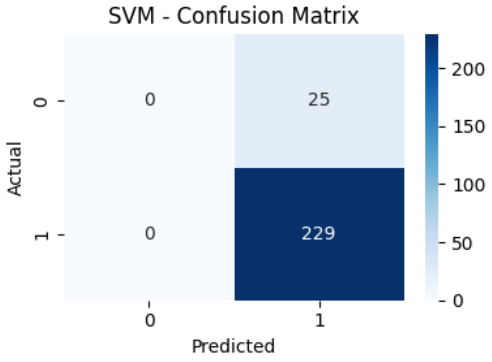
(a) KNN - Confusion matrix



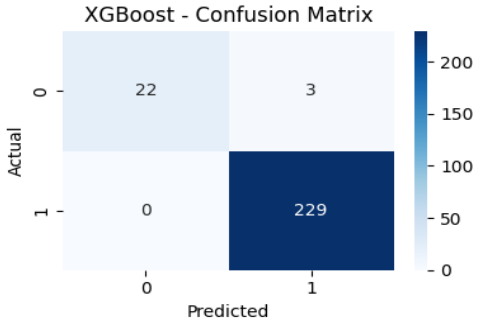
(b) LR – Confusion matrix



(c) RF – Confusion matrix



(c) SVM – Confusion matrix



(e) XGB – Confusion matrix

Figure 8. Models Confusion Matrix.

From Figure **8**, it can be seen that the True Positive (TP) demonstrated that the models successfully predicted numerous successful campaigns, this indicates a good recall and sensitivity. Also, the False Negatives (FN) appears to be that only few cases of successful campaigns being falsely predicted as unsuccessful. This means the models possess the capacity to minimize instances of missed positive instances. In addition, False Positives (FP) displays that the hardly ever fails to predict unsuccessful campaigns as successful. Lastly, the True Negatives (TN) shows that most unsuccessful campaigns were also accurately predicted. Furthermore, the Random Forest (RF) and XGBoost (XGB) models achieved

excellent performance and balanced the True Positive (TP) and the True Negative (TN) with nearly perfect classification of the matrix categories with the help of tuning. In addition, the False Positive (FP) and the False Negative (FN) indicates that the model make some few errors in identifying successful campaigns as successful campaigns that are unsuccessful.

Conclusion

In this research, extensive data preprocessing was conducted at the initial stage to ensure data quality and prevent multicollinearity problems before model construction. Five machine learning classifiers Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Extreme Gradient Boosting (XGB) were employed for accurate prediction of social media campaign post success. The classification performance indicated that LR, SVM, and KNN achieved good accuracy levels ranging from 85% to 89%, while the tree-based classifiers RF and XGB exhibited superior predictive capability with overall accuracies of 98% and 97%, respectively. These results clearly demonstrate that the chosen models are well-suited for predicting social media advocacy success.

While the findings of this study are demonstrated using the #JusticeForNavalAbbas campaign in Nigeria, the methodology and machine learning framework employed are generalizable to other social media advocacy campaigns, particularly those in African and other developing-country contexts that remain underrepresented in existing literature. The approach can be extended to any campaign dataset containing similar engagement features, offering a replicable framework for digital activism research globally.

Nonetheless, the study acknowledges certain limitations. These include the potential influence of pre-campaign activities particularly the initial broadcast of the campaign on the widely followed human rights radio program, Bereket Family which could not be empirically captured. Additionally, the concurrent use of engagement metrics as both predictive features and success label components represents a methodological consideration that future studies should address by incorporating temporally separated or pre-campaign features. Furthermore, the study is limited to a five-day data collection window across four platforms, which may not fully capture the complete lifecycle of the campaign. Future research should incorporate temporal features, sentiment analysis, and multi-campaign datasets spanning diverse sociopolitical contexts to further validate and strengthen the predictive framework introduced in this study.

Acknowledgement

We thank the International Conference on Advanced Science and Engineering (ICOASE 2025) for hosting this research.

Conflict of interest

None.

References

[1] G. Appel, L. Grewal, R. Hadi, and A. T. Stephen, "The future of social media in marketing," *J. Acad. Mark. Sci.*, vol. 48, no. 1, pp. 79–95, 2020, doi: 10.1007/s11747-019-00695-1.

[2] A. D. Oluwatoyosi, "Social Media Mobilization and Advocacy: A Case Study," *Journal of Information and Knowledge*, vol. 61, no. April, pp. 83–89, 2024, doi: 10.17821/srels/2024/v61i2/171307.

[3] E. Chukwuemeka, "The October 2020 #EndSARS Protest and the Transformation of the Nigerian State," *Journal of Political Science and International Relations*, no. March, 2023, doi: 10.11648/j.jpsir.20230601.13.

[4] O. OLUKEMI and K. Agber, "Social Media and Mass Mobilisation: An Analysis of X (Twitter) and the #EndSARS Protest," *New Frontiers: A Journal in the Humanities* , vol. 4, no. 1, pp. 30–47, Feb. 2026, Accessed: May 11, 2026. [Online]. Available: https://thesisprofs.org/index.php/new-frontiers/article/view/139

[5] S. A. Ahmad and A. B. Joseph, "Social Media and Women Mobilization for Community Development," *Journal of Media,Culture and Communication*, vol. 2, no. 36, pp. 39–49, 2023, doi: 10.55529/jmcc.36.39.49.

[6] D. Farsi, "Social media and health care, part i: Literature review of social media use by health care providers," *J. Med. Internet Res.*, vol. 23, no. 4, pp. 1–21, 2021, doi: 10.2196/23205.

[7] M. Kumari, "Social media and women empowerment," *International Journal of Scientific and Technology Research*, vol. 9, no. 3, pp. 626–629, 2020.

[8] V. Makatlal and D. Kumar, "The Impact of Social Media on Political Mobilization and Activism," *International Journal of Research Publication and Reviews Journal homepage: www.ijrpr.com*, vol. 3, no. 11, pp. 3468–3470, 2022, [Online]. Available: www.ijrpr.com

[9] M. I. Seelig, D. Millette, C. Zhou, and J. (Catherine) Huang, "A new culture of advocacy: An exploratory analysis of social activism on the web and social media," *Atl. J. Commun.*, vol. 27, no. 1, pp. 15–29, 2019, doi: 10.1080/15456870.2019.1540418.

[10] M. Tye, C. Leong, F. T. C. Tan, B. Tan, and Y. H. Khoo, "Social media for empowerment in social

movements: The case of Malaysia's grassroots activism," *Communications of the Association for Information Systems*, vol. 42, no. 1, pp. 408–430, 2018, doi: 10.17705/1CAIS.04215.

[11] S. Zilber, S. Monken, and F. Quevedo-Silva, "Adoption of social media by small- And medium-sized healthcare enterprises," *Brazilian Business Review*, vol. 16, no. 5, pp. 453–469, 2019, doi: 10.15728/bbr.2019.16.5.3.

[12] J. Stromer-Galley, F. Zhang, J. Hemsley, and S. Tanupabrungsun, "Tweeting the Attack: Predicting Gubernatorial Candidate Attack Messaging and Its Spread," *Int. J. Commun.*, vol. 12, pp. 3511–3532, 2018.

[13] I. A. Wulandari, M. Putriana, W. P. Sari, A. Soegiarto, and M. F. Rizki, "The Role of the Spotify Wrapped Digital Campaign on Spotify Brand Advocacy on Social Media Instagram, Twitter (X), and Tiktok," *Journal of Research in Social Science and Humanities*, vol. 4, no. 1, pp. 95–104, May 2024, doi: 10.47679/jrssh.v4i1.96.

[14] N. Cavus, A. S. Sani, Y. Haruna, and A. A. Lawan, "Efficacy of social networking sites for sustainable education in the era of COVID-19: A systematic review," *Sustainability (Switzerland)*, vol. 13, no. 2, pp. 1–18, Jan. 2021, doi: 10.3390/su13020808.

[15] A. Van Den Beemt, M. Thurlings, and M. Willems, "Towards an understanding of social media use in the classroom: a literature review," *Technology, Pedagogy and Education*, vol. 29, no. 1, pp. 35–55, 2020, doi: 10.1080/1475939X.2019.1695657.

[16] T. H. Dambo *et al.*, "Nigeria's #EndSARS movement and its implication on online protests in Africa's most populous country," *J. Public Aff.*, vol. 22, no. 3, 2022, doi: 10.1002/pa.2583.

[17] G. A. Pradhana and S. Tania, "Hyperreality of #blacklivesmatter movement on social media," *Jurnal Komunikasi: Malaysian Journal of Communication*, vol. 37, no. 3, pp. 288–303, 2021, doi: 10.17576/JKMJC-2021-3703-17.

[18] D. Buenaño-Fernández, D. Gil, and S. Luján-Mora, "Application of machine learning in predicting performance for computer engineering students: A case study," *Sustainability (Switzerland)*, vol. 11, no. 10, pp. 1–18, 2019, doi: 10.3390/su11102833.

[19] A. A. Lawan, N. Cavus, R. Yunusa, U. I. Abdulrazak, and S. Tahir, "Fundamentals of machine-learning modeling for behavioral screening and diagnosis of autism spectrum disorder," in *Neural engineering techniques for autism spectrum disorder. Volume 2, Diagnosis and clinical analysis*, 1st Editio., S. Jasjit and E.-B. Ayman S., Eds., London: Elsevier Academic Press, 2022, ch. Chapter 12, pp. 253–268. doi: 10.1016/B978-0-12-824421-0.00020-5.

[20] N. Cavus *et al.*, "A systematic literature review on the application of machine-learning models in behavioral assessment of autism spectrum disorder," *J. Pers. Med.*, vol. 11, no. 4, p. 299, Apr. 2021, doi: 10.3390/jpm11040299.

[21] H. Habehh and S. Gohel, "Machine Learning in Healthcare," *Curr. Genomics*, vol. 22, no. 4, pp. 291–300, 2021, doi: 10.2174/1389202922666210705124359.

[22] Jhaveri Siddharth, I. Khedkar, Y. Kantharia, and S. Jaswal, "Success prediction using random forest, catboost, xgboost and adaboost for kickstarter campaigns," *Proceedings of the 3rd International Conference on Computing Methodologies and Communication, ICCMC 2019*, no. Iccmc, pp. 1170–1173, 2019, doi: 10.1109/ICCMC.2019.8819828.

[23] I. Lee and Y. J. Shin, "Machine learning for enterprises: Applications, algorithm selection, and challenges," *Bus. Horiz.*, vol. 63, no. 2, pp. 157–170, 2020, doi: 10.1016/j.bushor.2019.10.005.

[24] N. Rahmayanti *et al.*, "ScienceDirect ScienceDirect Comparison of machine learning algorithms Comparison of machine learning algorithms to classify fetal health using cardiotocogram data to classify fetal health using cardiotocogram data," *Procedia Comput. Sci.*, vol. 197, no. 2021, pp. 162–171, 2022, doi: 10.1016/j.procs.2021.12.130.

[25] B. M. Muhammad, A. A. Lawan, J. Bala, T. S. Abdulrauf, and B. I. Muhammad, "Predictive modeling of student career pathways using machine learning techniques," *Journal of Statistical Sciences and Computational Intelligence*, vol. 1, no. 3, pp. 166–174, 2025, doi: 10.64497/jssci.95.

[26] F. Huang, J. Chen, Z. Lin, P. Kang, and Z. Yang, "Random forest exploiting post-related and user-related features for social media popularity prediction," *MM 2018 - Proceedings of the 2018 ACM Multimedia Conference*, pp. 2013–2017, 2018, doi: 10.1145/3240508.3266439.

[27] N. Bahrawi, "Sentiment Analysis Using Random Forest Algorithm-Online Social Media Based," *Journal of Information Technology and Its Utilization*, vol. 2, no. 2, p. 29, 2019, doi: 10.30818/jitu.2.2.2695.

[28] A. Gabay, A. Solomon, I. Guy, and B. Shapira, "Creating an Intelligent Social Media Campaign Decision-Support Method," *UMAP 2024 - Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*, pp. 149–158, 2024, doi: 10.1145/3627043.3659543.

[29] Z. Xu and M. Qian, "Predicting Popularity of Viral Content in Social Media through a Temporal-Spatial Cascade Convolutional Learning Framework," *Mathematics 2023, Vol. 11, Page 3059*, vol. 11, no. 14, p. 3059, Jul. 2023, doi: 10.3390/MATH11143059.

[30] M. Furini, F. Mandreoli, R. Martoglia, and M. Montangero, "A Predictive Method to Improve the Effectiveness of Twitter Communication in a Cultural Heritage Scenario," *ACM Journal on*

Computing and Cultural Heritage (JOCCH), vol. 15, no. 2, Apr. 2022, doi: 10.1145/3470786.

[31] D. T. Andariesta and M. Wasesa, "Machine learning models to predict the engagement level of Twitter posts: Indonesian e-commerce case study," *Procedia Comput. Sci.*, vol. 227, pp. 823–832, Jan. 2023, doi: 10.1016/J.PROCS.2023.10.588.

[32] I. Venkata and D. Srihith, "Predicting Success: The Impact of Machine Learning on Social Media Ad Classification," no. August, 2023, doi: 10.5281/zenodo.8214337.

[33] M. O. Biricik, Y. Eroğlu, and S. Mete, "Machine Learning-Powered Forecasting for Social Media Campaign Optimization in E-Commerce," *Journal of Information Analytics*, vol. 1, no. 1, p. 1, 2025.

[34] K. Vashisht, T. Kaur, S. Matlotia, A. Dhadwal, A. Sharma, and S. Shilpa, "Prediction Analysis on Trending Twitter Hashtags Using Machine Learning," *SSRN Electronic Journal*, no. January, pp. 0–6, 2024, doi: 10.2139/ssrn.4487088.

[35] S. K. Ghosh and F. Janan, "Prediction of Student ' s Performance Using Random Forest Classifier," pp. 7089–7100, 2021.

[36] B. Mahesh, "Machine Learning Algorithms - A Review," no. January 2019, 2020, doi: 10.21275/ART20203995.