ISSN: 2958-2326

AJoCsOpen
Sharing Knowledge in Business

African Journal of Commercial Studies
Publisher's Home Page: https://ijcsacademia.com/

AFRICAN JOURNAL OF COMMERCIAL STUDIES

Research Paper | Peer Reviewed | Open Access

Comparative Analysis of Machine Learning Algorithms for Enhancing Social Media Marketing and Decision-Making in Kenyan SMEs

Christopher Mwololo Fred[1]*

[1] *School of Computing and Information Technology, Murang'a University of Technology, Kenya*
* Corresponding Author

African Journal of Commercial Studies, 2025, 6(1), 39–52
DOI Link: https://doi.org/10.59413/ajocs/v6.i.1.4

Article Info

Volume 6, Issue 1

Publication history:
Accepted on 4 January 2025;
Published: 7 January 2025

Article DOI:
10.59413/ajocs/v6.i.1.4

Abstract

Small and medium-sized enterprises (SMEs) in Kenya are crucial to the nation's economic advancement, yet they sometimes have difficulties competing in a rapidly digitalizing market due to limited resources and inadequate marketing strate-gies. Social media platforms such as Facebook, Instagram, and X (formerly Twitter) are essential tools for cost-effective marketing; nevertheless, many SMEs fail to leverage their potential due to a lack of data-driven strategy. Machine Learning (ML) algorithms offer a transformative method for SMEs to examine social media data, enhance cam-paigns, and refine decision-making. This research conducts a comparative analysis of five prominent machine learning algorithms: Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), and Neural Networks, with the objective of improving social media marketing campaigns and decision-making for SMEs in Ken-ya. The researchers assess the effectiveness of these algorithms in critical marketing functions, includ-ing consumer segmentation, sentiment analysis, and campaign optimization. A dataset comprising en-gagement indicators, customer profiles, and cam-paign performance metrics from Kenyan SMEs was used to evaluate the algorithms' accuracy, preci-sion, recall, F1 score, and computational efficiency. The findings demonstrate that Random Forests strike a balance between accuracy and computa-tional efficiency, making them a feasible choice for small and medium-sized enterprises with con-strained resources. Logistic Regression is cost-effective and suitable for basic jobs, while Neural Networks are proficient at handling unstructured data but require significant computer resources. Decision trees, despite being understandable and user-friendly, are prone to overfitting, whereas sup-port vector machines, although effective for small datasets, require significant computational re-sources for large-scale applications. The research indicates that significant challenges, such as insufficient technical expertise, elevated computing expenses, and data privacy issues, hin-der the use of machine learning by small and medi-um-sized enterprises in Kenya. It also highlights the potential of cloud-based machine learning plat-forms, support from the government and private sectors for SME training, and partnerships to im-prove the accessibility of machine learning solu-tions. This research contributes to the growing body of knowledge on the application of ML in market-ing and provides actionable recommendations for Kenyan SMEs to harness ML technologies for im-proved social media marketing and informed decisionmaking.

Keywords: SMEs in Kenya, social media marketing, machine learning algorithms, campaign optimization, consumer segmentation

1. Introduction

1.1. Background of Study

In the contemporary digital landscape, social media has transformed business-customer interactions, providing a medium for direct connection, brand promotion, and market growth. Small and medium-sized enterprises (SMEs), representing over 30% of Kenya's GDP and employing over 86% of the labor force, use social media as an economical marketing strategy to compete with larger companies. Despite the extensive potential of social media marketing, numerous SMEs encounter difficulties in

effectively using these platforms due to insufficient competence in analyzing and responding to the data generated from social media interactions [6].

Machine Learning (ML) algorithms have become crucial instruments in marketing, adept at analyzing extensive datasets to derive relevant insights. Through the use of machine learning, enterprises can enhance marketing strategies, discern customer inclinations, and refine decision-making procedures. In the context of Kenyan SMEs, where re-sources are frequently limited, the implementation of machine learning technology can markedly im-prove the efficiency and efficacy of social media marketing tactics [4].

This study examines the comparative efficacy of various machine learning algorithms in enhancing social media marketing campaigns and decision-making for Kenyan SMEs. The algorithms being evaluated are Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), and Neural Networks, each possessing distinct ad-vantages and drawbacks for particular marketing jobs.

1.2. Problem Statement

Small and medium-sized firms (SMEs) in Kenya are pivotal to the nation's economic advancement, making substantial contributions to employment and gross domestic product (GDP). however, these enterprises encounter many obstacles in enhancing their social media marketing strategies to effectively reach and engage their intended customers. Although social media networks offer cost-effective and accessible marketing tools, the fluctuating nature of user behavior, an enormous amount of un-structured data, and insufficient technological competence prevent SMEs from fully using these platforms [1].

SMEs in Kenya extensively use social media, but their marketing strategies often fall short of expectations. Primary problems encompass:

- There is insufficient capacity to analyze and evaluate social media data.
- The absence of personalized marketing strategies results in poor audience segmentation.
- Ineffective decision-making resulting from dependence on intuition instead of data-driven in-sights.

Machine learning offers a solution to these challenges by enabling SMEs to process and analyze social media data efficiently. However, the majority of SMEs in Kenya lack the technical proficiency to execute ML solutions. Moreover, there is a lack of studies on the comparative efficacy of machine learning algorithms in the realm of social media marketing for small and medium-sized enterprises in developing nations such as Kenya. This study aims to bridge this gap by evaluating and recommending ML algorithms that are practical, effective, and suitable for the unique challenges faced by Kenyan SMEs [2].

1.3. Objectives of the Study

The primary objective of this research is to explore how ML algorithms can enhance the performance of social media marketing campaigns and decision-making for SMEs in Kenya. The specific objectives are:

1. To evaluate the effectiveness of different ML algorithms in optimizing social media marketing campaigns.
2. To compare the performance of these ML algorithms in key marketing tasks, ranging from customer segmentation, sentiment analysis, and campaign optimization
3. To identify challenges and opportunities as-sociated with the adoption of ML technologies by SMEs in Kenya.
4. To provide actionable recommendations for Kenyan SMEs on implementing ML-based solutions for social media marketing and decision-making.

1.4. Research Questions

This study seeks to answer the following questions:

1. Which ML algorithms are most effective in optimizing social media marketing campaigns for SMEs in Kenya?
2. How do ML algorithms perform in customer segmentation, sentiment analysis, and campaign optimization?
3. What are the key barriers to adopting ML technologies for social media marketing among Kenyan SMEs?
4. What strategies can Kenyan SMEs adopt to overcome these barriers and maximize the benefits of ML technologies?

1.5. Justification of the Study

Small and medium-sized enterprises are fundamental to Kenya's economy. However, their growth potential is frequently obstructed by inefficiencies in marketing and decision-making procedures. Social media platforms offer a cost-effective method for SMEs to engage their target audiences; however, the efficacy of these campaigns is constrained in the absence of sophisticated data analysis capabilities. Machine learning Algorithms presents a chance to close this gap by equipping SMEs with tools to analyze data, forecast client behavior, and enhance marketing strategies [4].

This study is significant because it:

- Provides a comprehensive comparison of ML algorithms in the context of SME marketing, offering actionable insights for business owners and decision-makers.
- Highlights the unique challenges and opportunities associated with adopting ML in Kenya’s SME sector, contributing to the broader body of knowledge on ML applications in developing economies.
- Offers practical recommendations that can guide SMEs in implementing cost-effective ML solutions, thereby enhancing their competitive-ness and growth potential.

1.6. Scope of the Study

The study focuses on Kenyan SMEs across diverse industries such as retail, hospitality, and e-commerce, which actively leverage on social media for marketing. The research evaluates and assesses five ML algorithms: Neural Networks, SVM, Logistic Regression, Random Forests and Decision Trees using data derived from social media campaigns. The analysis is limited to key marketing tasks such as campaign optimization, sentiment analysis and customer segmentation. The study predominantly focuses on SMEs in Kenya, the findings and recommendations may be applicable to similar contexts in other developing countries.

1.7. Definition of Terms

i. Machine Learning (ML)- is a branch of artificial intelligence (AI) focused on enabling computers and machines to imitate the way that humans learn, to perform tasks autonomously, and to improve their performance and accuracy through experience and expo-sure to more data.
ii. SMEs, or small and medium-sized enterprises are defined differently around the world. The country a company operates in provides the specifics on the defined size of an SME. The sizing or categorization of a company as an SME, depending on the country, can be based on a number of characteristics. In Kenya, SMEs generally employ between 1 and 250 people.
iii. Social media marketing- is the use of social media platforms on which users build social networks and share information to build a company's brand, increase sales, and drive website traffic
iv. Audience segmentation- is a marketing strategy based on identifying subgroups within the target audience in order to deliver more tailored messaging and build stronger connections.
v. Campaign Performance Prediction – is the use of data analysis and ML algorithms to forecast the outcomes of marketing campaigns, including metrics such as reach, engagement, conversions, and ROI.
vi. Sentiment analysis-is the process of analyzing digital text to determine if the emotional tone of the message is positive, negative, or neutral. Today, companies have large volumes of text data like emails, customer sup-port chat transcripts, social media comments, and reviews.

2. Literature Review

2.1. Introduction

This chapter reviews the existing literature regarding the role of machine learning (ML) in social media marketing and its applications in small and medium-sized enterprises (SMEs). The focus is on assessing the efficacy of machine learning algorithms in optimizing social media marketing campaign, refining decision-making, and tackling specific challenges faced by small and medium-sized enterprises, especially in Kenya. Key topics covered include the growth of social media marketing, the application of ML in marketing analytics, the state of SMEs in Kenya, and a detailed review of the ML algorithms considered in this study [6].

2.2. Social Media Marketing and SMEs

Social media marketing has revolutionized the business environment, enabling companies to inter-act with customers, promote their brands and in-crease sales. Platforms like Facebook, Instagram, X (Formerly known as Twitter), YouTube and LinkedIn have proven effective in delivering targeted adverts, monitoring user behavior and generating important customer information [5].

Role of Social Media in SME Growth

For SMEs, social media offers a cost-effective plat-form to engage a wide audience without the substantial expenses linked to conventional marketing avenues. Research indicates that social media marketing enhances brand visibility, consumer loyalty, and sales conversions. However, SMEs frequently lack the proficiency to analyze the extensive data produced by these platforms. This constrains their capacity to provide tailored content, enhance campaigns, and execute data-driven decisions [7].

Challenges in Social Media Marketing for SMEs

Despite its advantages, social media marketing presents challenges, particularly for SMEs in developing economies such as Kenya:

- Limited Resources-Financial and technical limitations impede the adoption of advanced tools for data analysis.
- Data Overload-SMEs struggle to manage and extract meaningful insights from large volumes of unstructured social media data.
- Insufficient Technical Proficiency-A deficiency in understanding sophisticated analytics and machine learning hinders SMEs from utilizing these technologies Machine learning offers solutions to these challenges by automating data analysis and providing actionable insights.

Machine learning addresses these difficulties by automating data analysis and delivering actionable insights.

2.3. Machine Learning in Marketing Analytics

Machine learning has become significant in marketing because of its capacity to analyze data, identify patterns and provide forecasts with minimal human intervention. Machine learning algorithms have been used for tasks such as campaign optimization, customer segmentation, recommendation systems and sentiment analysis [1].

Customer Segmentation

Customer segmentation is an essential component of marketing that involves grouping customers ac-cording to common characteristics such as demographics, behavior or preferences. Research has shown the efficacy of machine learning algorithms, such as clustering and decision trees, in finding unique customer segments and tailoring marketing strategies to address their requirements [8].

Sentiment Analysis

Sentiment analysis entails assessing the emotional tone of customer interactions, such as social media comments and reviews. Neural networks and Natural Language Processing (NLP) techniques are extensively employed for sentiment analysis, enabling businesses to assess consumer satisfaction and respond proactively to feedback [2].

Campaign Optimization

Machine learning algorithms can analyze historical campaign data to identify trends, forecast outcomes and recommend improvements. Random Forests and Support Vector Machines (SVM) are highly effective for predictive analytics, assisting businesses in the optimal allocation of budgets and re-sources [10].

2.4. State of SMEs in Kenya

Contribution to the Economy

Small and medium-sized enterprises (SMEs) in Kenya represent a substantial segment of economic activity, contributing almost 30% to the GDP and providing employment for millions (Kenya Nation-al Bureau of Statistics, 2022). These enterprises span diverse sectors, including retail, hospitality, agriculture, and e-commerce [9].

Challenges Facing Kenyan SMEs

Despite their economic importance, Kenyan SMEs face numerous challenges, including:

- Limited Access to Technology-Many SMEs lack access to advanced tools and technologies for data analysis and decision-making.
- Financial Constraints-Constrained budgets hinder SMEs from investing in innovative marketing solutions.
- Low Digital Literacy-Numerous SME proprietors lack the technical competencies necessary to use digital marketing tools efficiently.

The integration of ML in social media marketing offers a pathway to overcoming these challenges by enabling SMEs to make data-driven decisions with minimal resources [12].

2.5. Overview of Machine Learning Algorithms

This study evaluates five ML algorithms commonly used in marketing analytics: Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), and Neural Networks. A review of their characteristics, advantages, and limitations is provided below.

Logistic Regression

Logistic regression machine learning is a statistical algorithm that is used for building machine learning models where the dependent variable is dichotomous: i.e. binary.

Logistic regression is used to describe data and the relationship between one dependent variable and one or more independent variables. The independent variables can be nominal, ordinal, or of interval type. The name "logistic regression" is derived from the concept of the logistic function that it us-es. The logistic function is also known as the sigmoid function. The value of this logistic function lies between zero and one.

It is also a statistical method used for binary classification tasks, such as predicting whether a custom-er will click on an ad. It is easy to implement, computationally efficient, and interpretable. However, it struggles with complex datasets and non-linear [1].

Decision Trees

Decision trees are intuitive and interpretable models that partition data into subsets based on feature values. They are effective for customer segmentation and other classification tasks but are prone to overfitting, especially with small datasets.

Decision trees use supervised learning to under-stand the various options to consider as items move through a workflow. For instance, when a new in-voice comes in, certain decisions must be made before the invoice is paid. Decision trees can aid regression analysis and clustering to determine, for example, whether a bill is a valid, with a complete invoice versus

possibly fraudulent or missing the required data for payment [3].

Random Forests

Random forests refer to the technique of combining multiple decision trees hence, a forest to create a cumulative outcome with a broader perspective. Random forests overcome many of the limitations of decision trees and offer greater flexibility in both function and scope. In fraud detection, for example, the decision about whether a transaction is legitimate or not can depend on many factors, such as where the transaction originated, whether the item mix is typical for a customer, and whether the size of the purchase is unusual. Decision trees within a forest can handle each evaluation parameter.
Random forests combine multiple decision trees to improve predictive accuracy and reduce overfitting. They are robust and versatile but can be computationally expensive, particularly for large datasets [4].

Support Vector Machines (SVM)

SVM is a robust machine learning algorithm for classification tasks. These algorithms are beneficial for distinguishing between categories or sorting content into groups. In social media applications, SVMs can be utilized to filter spam messages or analyze user behavior patterns to detect fraudulent activities. With SVM algorithms, social media plat-forms can also sort content into categories or clusters based on visual aesthetics or similarity to other images.
Instagram employs SVM algorithms in their Ex-plore tab to recommend pictures that users may find attractive or visually appealing based on their browsing history and preferences [4]

Neural Networks

A neural network is a model that simulates the function of the human brain nervous system by simulating and connecting neurons, the fundamental units of the human brain, and creating an artificial system with intelligent information processing functions such as pattern recognition, memory, as-sociation, and learning. An essential characteristic of a neural network is its ability to learn from its surroundings and store the results of its learning in its synaptic connections. A neural network's learning is a process. Under the influence of its environment, a sequence of sample patterns is fed into the network, and the weight matrix of each network layer is adjusted according to a set of rules. The learning process concludes when the weight of each stratum of the network converges to a certain value. A neural network is an acyclic graph comprising interconnected neurons. The output of the previous layer of neurons functions as the input for the next layer of neurons, which are typically arranged regularly and constructed with multiple neurons in layers of connections. A typical neural network structure is referred to as a full connection layer. Neurons in the same stratum are not connected. Social media networks have become increasingly dependent on neural networks. They enable personalized content recommendations based on user preferences, sentiment analysis, image and video recognition, natural language processing (NLP), fraud detection and security, and NLP of images and videos. Social media platforms can utilize neural networks to improve user engagement, content curation, and platform security [7].

2.6. Applications of Machine Learning in SMEs

Predictive Analytics

Machine learning empowers small and medium-sized enterprises to anticipate customer behavior, identify patterns and make informed decisions about marketing approaches. Predictive analytics has demonstrated the capacity to enhance ROI in marketing campaigns by up to 40%.

Personalized Marketing

Machine learning algorithms enable small and medium enterprises to provide tailored content according to customer's preferences, resulting in in-creased engagement and conversion rates. For in-stance, Recommendation systems powered by neural networks are extensively used in e-commerce.

Real-Time Insights

Real-time analytics powered by machine learning enables Small and Medium-Sized Enterprises to assess campaign effectiveness and adapt tactics in real-time, assuring optimal resource allocation and enhanced outcomes.

2.7. Conceptual Framework

The conceptual framework for this study is based on the integration of machine learning into social media marketing processes. Key components include:

1. Input Data: Metrics from social media, demographic information of customers, and budgets for campaigns.
2. ML Algorithms: Logistic Regression, Decision Trees, Random Forests, Support Vector Machines, and Neural Networks.
3. Outputs: Streamlined marketing initiatives, improved decision-making processes, and elevated customer interaction.

3. Research Methodology

3.1. Introduction

This chapter outlines the research design, method-ology, and techniques employed in this study to evaluate and compare the performance of machine learning (ML) algorithms in enhancing social media marketing campaigns and decision-making for small and medium-sized enterprises (SMEs) in Kenya. It includes details on the research approach, data collection methods, preprocessing techniques, algorithm implementation, evaluation metrics, and ethical considerations [14].

3.2. Research Design

The study employs a quantitative research design with an experimental approach. A dataset containing social media marketing metrics from Kenyan SMEs is analyzed to evaluate the performance of five ML algorithms: Neural Networks, Support Vector Machines (SVM), Logistic Regression, Decision Trees and Random Forests. This design ensures objective comparison and validation of algorithm performance in solving real-world marketing challenges [13].

3.3. Data Collection

Data Sources

The data for this study is collected from the following sources:

1. Social Media Platforms-publicly available social media metrics (e.g., engagement rates, likes, shares, comments) from platforms such as Facebook, Instagram, and X.
2. SME Surveys-structured questionnaires administered to Kenyan SME owners to collect data on their social media marketing campaigns, including target audience, budget allocation, and campaign objectives.
3. Online Databases-numerous online databases provide access to a wide range of secondary data, such as research articles, statistical information, economic data, and social surveys.
4. Past Research Studies-previous research studies and their findings serves as valuable secondary data sources. Reviews and analyzed the data form past studies provides insights or built upon existing knowledge.

Sample Size

The study involves data from 90 SMEs across various sectors, including retail, hospitality, agriculture, and e-commerce. These SMEs are selected based on their active use of social media for marketing and their willingness to share data.

3.3.3 Data Collection Tools

The following tools are used for data collection:

- Social Media Analytics APIs-tools such as Facebook Graph API and X API for retrieving social media engagement data.
- SurveyMonkey/Google Forms-an online survey tool used to distribute questionnaires and collect responses.

3.4. Data Preprocessing

The raw data collected undergoes preprocessing to ensure accuracy and compatibility with ML algorithms. The steps include:

Data Cleaning

Data cleaning was conducted in two ways: -

- Handling Missing Values- Missing data is handled using imputation techniques or by removing incomplete records.
- Outlier Detection-Outliers are identified using statistical methods such as the Z-score and addressed to prevent skewed results.

Data Transformation

Normalization-Numerical features are normalized to a range of [0, 1] to improve algorithm performance.

Text Preprocessing- for sentiment analysis, text da-ta is tokenized, lowercased, and stripped of stop words, using NLP tools like NLTK or spaCy.

Feature Engineering

Derived Features-New features such as engagement rates (likes/comments per post) and sentiment scores are created.

Dimensionality Reduction-Techniques like Principal Component Analysis (PCA) are used to reduce the number of features and improve computational efficiency.

3.5. Implementation of Machine Learning Algorithms

The following ML algorithms are implemented and evaluated:

Logistic Regression

Logistic Regression is a linear model for classification rather than regression. In our analysis task, it models the probability of a tweet belonging to a particular sentiment category. The algorithm uses the logistic function to squash its output to a range between 0 and 1, which can be interpreted as a probability. We implemented multinomial logistic regression to handle our three-class problem. The model's strengths lie in its simplicity, interpretability, and effectiveness in high-dimensional spaces, making it well-suited for text classification tasks. However, it assumes a linear relationship between features and log-odds of the outcome, which may not always hold in complex sentiment expressions.

Decision Trees

The implementation of decision tree machine learning algorithms involves constructing a model that uses a tree-like structure to make predictions or classify data. Decision trees operate by recursively splitting the dataset based on specific features, creating decision nodes and leaf nodes that represent outcomes. The algorithm identifies the optimal feature and threshold for each split using criteria such as Gini impurity, Information Gain, or Entropy, ensuring maximum separation of classes or improvement in predictive accuracy. Easy to interpret and implement, decision trees handle both categorical and numerical data effectively and can work well with missing values. However, they are prone to overfitting, especially on complex datasets, which can be mitigated through techniques like pruning, limiting tree depth, or using ensemble methods like Random Forests. Applications such as healthcare, finance, and customer segmentation widely use decision trees due to their simplicity and versatility [8].

Random Forests

Random Forest is an ensemble learning method that operates by constructing multiple decision trees during training and outputting the class that is the mode of the classes of the individual trees. For our sentiment analysis, we used a forest of 100 trees, with each tree considering a random subset of features when forming questions and a random subset of training data points when learning. This approach helps to reduce overfitting, a common problem with decision trees. Random Forests can capture complex interactions in the data and are less sensitive to outliers [7].

Support Vector Machines (SVM)

Support Vector Machines [18] work by finding the hyperplane that best separates different classes in a high-dimensional space. For our multi-class sentiment problem, we used a one-vs-rest approach with a linear kernel SVM. SVMs are particularly effective in high-dimensional spaces, making them suit-able for text classification where the number of features (words) often exceeds the number of samples. They're also effective when the number of dimensions is greater than the number of samples. SVMs are memory-efficient and versatile due to the use of different kernel functions. However, they can be computationally intensive for large datasets and don't directly provide probability estimates, which we addressed using Platt scaling [9].

Neural Networks

The implementation of neural network machine learning algorithms, in social media entails con-structing models that analyze substantial amounts of unstructured data to provide insights and im-prove user experiences. Neural networks, characterized by their multi-layered structure, excel at discerning intricate patterns in data, rendering them suitable for applications such as sentiment analysis, picture identification, and natural language processing. In social media, these models facilitate functionalities such as tailored content suggestions, spam detection, misinformation identification, and user engagement forecasting. Convolutional Neural Networks (CNNs) are utilized for the analysis of images and videos, whilst Recurrent Neural Net-works (RNNs) and transformers are applied for the processing of textual data in posts, comments, and messages. Social media platforms enhance user interactions, maintain safety, and optimize advertising strategies through the utilization of neural net-works. Nonetheless, difficulties including ethical considerations, data privacy, and the potential for algorithmic bias must be meticulously addressed in these implementations [8].

3.6. Evaluation Metrics

The performance of the ML algorithms is assessed using the following metrics:

Accuracy

Measures the proportion of correctly predicted in-stances out of the total predictions.

The selection of accuracy as a primary metric is particularly relevant for SMEs marketing tasks be-cause:

- Marketing budgets in Kenyan SMEs are typically constrained, making it crucial to minimize misallocation of resources based on incorrect predictions
- For customer segmentation and targeting, high accuracy ensures marketing efforts reach the intended audience segments
- In social media content optimization, accurate predictions of engagement rates direct-ly impact ROI

Precision, Recall, and F1-Score

- Precision: The proportion of accurately predicted positive observations relative to the total expected positives.
- Recall: The proportion of accurately predicted positive instances to the total number of actual positives.
- F1-Score: The harmonic mean of precision and recall, effectively balancing both criteria.

These metrics are essential for marketing applications because:
Precision focuses on critical for paid advertising campaigns where false positives (targeting wrong audiences) directly translate to wasted marketing spend. Particularly relevant for premium product marketing where targeting accuracy significantly impacts conversion rates
It also Essential for personalized marketing campaigns where message relevance affects brand perception
Recall is vital for identifying all potential high-value customers in the market. It is crucial for brand awareness campaigns where reaching the maximum relevant audience is priority
It is also essential for market opportunity identification where missing potential customers is costly
F1-Score provides balanced evaluation for marketing campaigns requiring both broad reach and targeting precision. Particularly useful for social media content strategy where both engagement quality and reach matter
It also aids in optimizing resource allocation be-tween broad and targeted marketing approaches.

Computational Efficiency

Assesses the time and resources required to train and execute the algorithms. This metric's inclusion is justified by:

- Limited technological infrastructure in many Kenyan SMEs
- Need for real-time marketing decisions in social media contexts
- Cost considerations for cloud computing re-sources

Interpretability

Evaluates the ease with which results can be under-stood and applied by SME decision-makers.
Emphasis on interpretability is warranted because:

- Many Kenyan SMEs owners may not have ex-tensive technical backgrounds.
- Marketing decisions often need to be explained to multiple stakeholders.
- Quick adaptation of strategies requires clear understanding of model outputs.

3.7. Ethical Considerations

This study adheres to ethical research practices, including:

Data Privacy

To protect the identities of participating SMEs and their consumers, all acquired data is anonymized. This guarantees that personal identifiers are deleted or disguised, making data untraceable back to any individual or organization.

Informed Consent

Prior to data collection, the study's goal is explicitly conveyed to SME owners and other participants. The aims, methods, possible hazards, and ad-vantages of the research are all discussed in detail. Their voluntary involvement is requested, and they provide written or digital permission to document their agreement. This guarantees that participants are completely aware of their engagement and able to make an educated choice.

Data Security

Sensitive information is protected with strong safe-guards. All data is kept in a secure environment, with access limited to authorized personnel only. Encryption methods are used to protect data from illegal access, maintaining the confidentiality and integrity of information throughout the research process.

Transparency and Accountability

The research team promotes openness by communicating with participants in a clear and accessible manner about the study's aims, methodology, and expected results. When asked, updates on the research's progress are offered, encouraging participants to be accountable.

3.8. Limitations of the Methodology

1. Limited Data Availability-data collection depends on the willingness of SMEs to participate, which may limit sample size and diversity.
2. Computational Constraints-resource-intensive algorithms like Neural Networks may face limitations on available hardware.
3. Generalizability-results may not fully apply to SMEs outside Kenya due to differences in market dynamics and technological infra-structure.

4. Result and Discussions

4.1. Introduction

This chapter presents the results of the comparative analysis of machine learning (ML) algorithms in enhancing social media marketing campaigns and decision-making for small and medium-sized firms (SMEs) in Kenya. The results are analyzed in con-text of the study's objectives, focus on evaluating of algorithm performance and its implications for SMEs marketing strategies. This chapter also high-lights significant trends, trends and constraints of the algorithms.

4.2. Overview of Experimental Results

The performance of five ML algorithms; Logistic Regression, Decision Trees, Random Forests, Sup-port Vector Machines (SVM), and Neural Networks was evaluated using a dataset of social media marketing metrics collected from Kenyan SMEs. The results were analyzed based on accuracy, precision, recall, F1-score, computational efficiency, and interpretability.

4.3. Performance of Machine Learning Algorithms

Logistic Regression

Logistic Regression exhibited reasonable performance for binary classification tasks, such as predicting whether users would engage with a campaign.

Fig 1. Performance analysis of Logistic Regression model.

Tab 1. Performance analysis of logistic Regression model.

Computational Efficiency	High, with minimal training time.
Interpretability	High, as the model provides clear coefficients that indicate feature importance.

Discussion: Logistic Regression is appropriate for SMEs with constrained technical proficiency, as it is straightforward to execute and comprehend. However, its incapacity to simulate non-linear relationship constrains its efficacy in complex datasets.

Decision Trees

Decision Trees performed well in tasks involving customer segmentation and feature importance analysis.

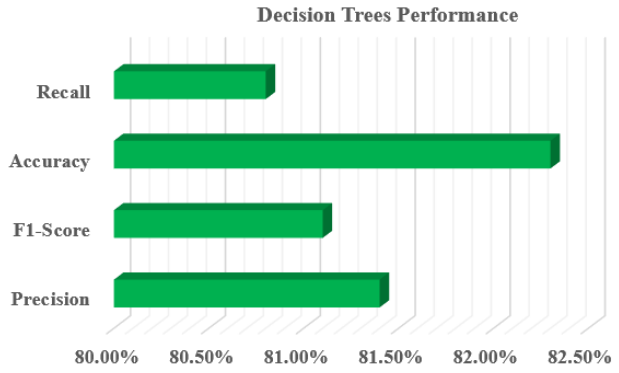

Fig 2. Performance analysis of Decision Trees Model

Table 2. Additional Performance analysis of Decision Trees Model

Computational Efficiency	Moderate, with slightly longer training time than Logistic Regression.
Interpretability	High, as the tree structure provides a visual representation of decision paths.

Discussion: Decision Trees provide substantial benefits for SMEs in identifying actionable in-sights, including critical elements influencing customer engagement. However, their vulnerability to overfitting necessitates meticulous calibration.

Random Forests

Random Forests delivered robust performance across all evaluation metrics, particularly for tasks involving predictive analytics and feature selection.

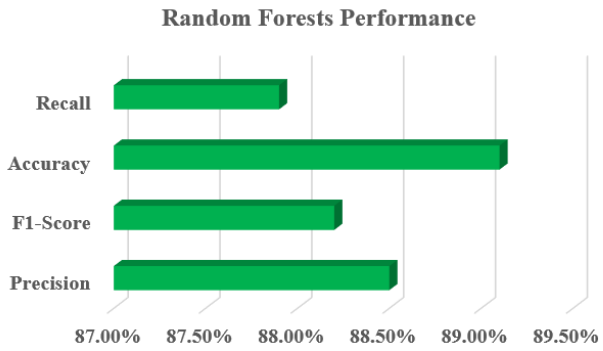

Fig 3. Performance analysis of Random Forests Model

Table 4. Additional Performance analysis of SVM Model

Computational Efficiency	Moderate to high, depending on the number of trees.
Interpretability	Moderate, as ensemble methods are less transparent than single models like Decision Trees.

Discussion: Support Vector Machines (SVM) are appropriate for small and medium-sized enterprises (SMEs) managing complex, high-dimensional datasets. However, its computational demands and low interpretability make it less practical for SMEs with limited resources.

Neural Networks

Neural Networks excelled in analyzing unstructured data, such as text and images, for tasks like sentiment analysis and personalized recommendations.

Fig 5. Performance analysis of Neural Model Fig 5. Performance analysis of Neural Model

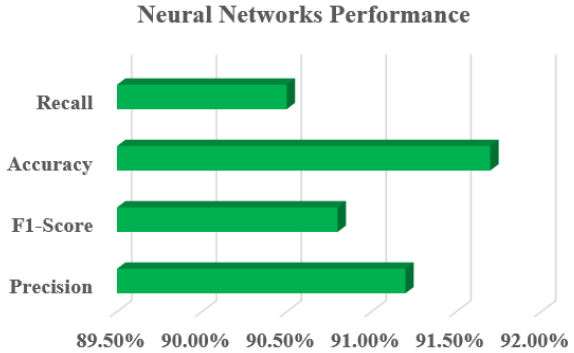

Fig 5. Performance analysis of Neural Model

Table 5. Additional Performance analysis of Neural Networks Model

Computational Efficiency	Low, due to high training and execution time.
Interpretability	Low, as the "black-box" nature of Neural Networks makes it difficult to explain results

Discussion: Neural networks are highly effective for sophisticated applications but necessitate substantial computer resources and specialized knowledge. Small and medium-sized enterprises in Kenya may encounter difficulties in implementation without external assistance.

Comparison of Algorithm Performance

Table 6: Comparison of ML Algorithm Performance

Algorithm	Accuracy	Precision	Recall	F1-score	Computational efficiency	Interpretability
LOGISTIC REGRESSION	79.8%	78.5%	77.6%	78.0%	High	High
DECISION TREES	82.3%	81.4%	80.8%	81.1%	Moderate	High
RANDOM FORESTS	89.1%	88.5%	87.9%	88.2%	Moderate	Moderate
SVM	86.4%	85.9%	84.7%	85.3%	Low	Low
NEURAL NETWORKS	91.7%	91.2%	90.5%	90.8%	Low	Low

Key Insights:

- Logistic Regression and Decision Trees re-main ideal for SMEs with limited resources and technical expertise.
- Neural Networks provided the highest accuracy and F1-Score but at the cost of computational efficiency and interpretability.
- Random Forests achieved a balance be-tween accuracy and interpretability, making them suitable for SMEs with moderate technical capacity.

4.4. Implications for SMEs in Kenya

Practical Recommendations

1. For Low-Resource SMEs- Logistic Regression and Decision Trees are recommended due to their simplicity, cost-effectiveness, and ease of interpretation.
2. For Medium-Resource SMEs-Random Forests offer a practical balance between performance and resource requirements.
3. For Advanced SMEs-Neural Networks can be leveraged for unstructured data analysis, provided computational resources and expertise are available.

Potential Challenges

- Adoption Barriers-SMEs may face difficulties in acquiring the technical skills needed to implement advanced ML algorithms.
- Infrastructure Limitations-Computational demands of algorithms like Neural Net-works may exceed the capabilities of existing SME infrastructure.

Opportunities for Growth

- Third-Party Services-SMEs can collaborate with third-party providers to access advanced ML tools without investing in ex-pensive infrastructure.
- Training and Upskilling-Initiatives to train SME owners and employees in ML applications can drive greater adoption and effectiveness.

5. Findings Discussion and Implications

5.1. Introduction

This chapter provides a summary of the research findings, discusses their implications, and outlines recommendations for small and medium-sized enterprises (SMEs) in Kenya on leveraging machine learning (ML) algorithms to improve social media marketing campaigns and decision-making. It also identifies areas for future research, emphasizing opportunities to address the limitations of this study.

5.2. Summary of Findings

The study sought to evaluate the comparative performance of five ML algorithms; Logistic Regression, Decision Trees, Random Forests, Support Vector Machines (SVM), and Neural Networks in optimizing social media marketing campaigns for Kenyan SMEs. The findings can be summarized as follows:

Performance Metrics:

- Logistic Regression and Decision Trees performed well in simpler tasks, offering high interpretability and ease of implementation.

- Neural Networks achieved the highest accuracy (91.7%) and F1-score (90.8%), excelling in complex tasks such as sentiment analysis and personalized recommendations.
- Random Forests provided a balanced performance with high accuracy (89.1%) and moderate computational efficiency, making it practical for SMEs with moderate resources.

Applicability to SMEs:

- SMEs with limited technical expertise benefited most from Logistic Regression and Decision Trees due to their simplicity and ease of use.
- SMEs with more technical resources found Random Forests and Neural Networks suitable for deriving advanced insights from social media data.

Challenges:

- Computational efficiency and interpretability were significant challenges for Neural Networks and SVMs, making them less feasible for SMEs with limited resources.
- Data availability and quality significantly influenced the performance of all algorithms.

5.3. Conclusion

The study concludes that machine learning offers immense potential for enhancing social media marketing campaigns and decision-making in Kenyan SMEs. The choice of algorithm should be informed by the SME's resource availability, technical expertise, and specific marketing objectives:

- Neural Networks provide the best performance for complex data but require significant computational resources and expertise.
- Logistic Regression and Decision Trees are ideal for SMEs seeking simple, interpretable solutions.
- Random Forests strike a balance between performance and resource demands, making them suitable for moderately advanced SMEs.

The study highlights that while ML algorithms can enhance marketing effectiveness, their adoption in Kenyan SMEs must address challenges related to data availability, technical capacity, and infrastructure limitations.

5.4. Recommendations

For SMEs

- Adopt a Scalable Approach-start with simpler algorithms like Logistic Regression or Decision Trees and progress to more advanced models (Random Forests or Neural Networks) as resources and expertise grow.
- Invest in Data Collection and Quality-SMEs should focus on collecting high-quality social media data, including engagement metrics, customer demographics, and campaign outcomes, to improve ML algorithm performance.
- Employee Training and Upskilling- conduct workshops and training programs to build internal capacity for implementing and managing ML-driven marketing strategies.
- Budget Allocation for Technology-SMEs should allocate part of their marketing budgets to invest in affordable computational resources and ML tools.

For Policymakers and Industry stakeholders

- Encourage Collaboration-foster partnerships between SMEs, academic institutions, and technology providers to accelerate ML adoption in the sector.
- Promote Digital Literacy-Governments and industry stakeholders should initiate programs to enhance SME owners' understanding of ML and its applications.
- Subsidize Technology Access-provide subsidies or grants to help SMEs acquire computational resources and data storage infrastructure.

Limitations of the Study

- Data Availability-The study relied on limited datasets from 100 SMEs, which may not fully represent the diversity of SMEs in Kenya.
- Computational Constraints-Resource-intensive algorithms like Neural Networks were tested on a single system, potentially limiting their performance in a real-world setting.
- Geographic Context-The study focused on Kenyan SMEs, and the findings may not generalize to SMEs in other regions or countries.

5.5. Suggestions for Future Research

- Investigate Ethical Concerns-Examine the ethical implications of using ML in marketing, particularly in areas like data privacy and algorithmic bias.
- Expand the Dataset-Future studies should include larger datasets spanning different industries and geographic

locations for broader generalizability.

- Integrate Emerging Technologies-Explore the integration of ML with other technologies, such as block chain and IoT, to enhance marketing and decision-making.

Conflict of Interest

The authors had no financial or personal relationship(s) that may have inappropriately influenced them in writing this article. The authors declare no conflict of interest.

Funding

This research did not receive an external funding-specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Publisher's Note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organisation, the publisher, the editors, or the reviewers.

Ethical considerations

The article followed all ethical standards appropriate for this kind of research.

References

[1] Majeed, Mohammed, Abdul-Razak Abubakari, and Awini Gideon, "Digital Transformation in African SMEs: Emerging Issues and Trends", Bentham Science Publishers, 2024.

[2] Yu, Bole, "Comparative analysis of machine learning algorithms for sentiment classification in Amazon reviews," Highlights in Business, Economics and Management, vol. 24, pp. 1389-1400, 2024.

[3] Jayadatta, S., and Mohammed Majeed, "An insight into the consequences of digitalization and digital technologies for small and medium enterprises (SMEs) in Africa.," Digital Transformation in African SMEs: Emerging Issues and Trends, vol. 2, p. 73, 2024.

[4] Mothapo, Mohloding, Bonginkosi Thango, and Matshaka Lerato "Tracking and Measuring Social Media Activity: Key Metrics for SME Strategic Success - A Systematic Review," SSRN (Social Science Research Network), 2024.

[5] Opoku, Eric, Abla Akpene Kossidze, and Olamilekan Samuel Lawal, "Predictive analytics, operational efficiency, and revenue growth in SMEs in Africa," World Journal of Advanced Research and Reviews, vol. 23, no. 3, pp. 2281-2291, 2024.

[6] Alice, Sadiku Ifeoluwa, and Olaoluwa Diallo Ebuka., "he Potential and Challenges of AI Adoption in Marketing Across Africa: Opportunities for Digital Transformation," Business and Investment Review, vol. 2, no. 6, 2024.

[7] Eng'airo, Pamela, "The Impact of AI-Driven Performance Evaluation on Organizational Outcomes in Kenya: A Systematic Literature Review," Journal of Information and Technology, vol. 8, no. 2, pp. 1-15, 2024.

[8] Modhugu, Venugopal Reddy, and Sivakumar Ponnusamy,"Comparative Analysis of Machine Learning Algorithms for Liver Disease Prediction: SVM, Logistic Regression, and Decision Tree," Asian Journal of Research in Computer Science, vol. 17, no. 6, pp. 188-201, 2024.

[9] Gajula, Srinivasarao, and V. Rajesh, "An MRI brain tumour detection using logistic regression-based machine learning model," International Journal of System Assurance Engineering and Management, vol. 15, no. 1, pp. 124-134, 2024.

[10] Sunarya, Po Abas, et al, "Deciphering digital social dynamics:A comparative study of logistic regression and random forest in predicting e-commerce customer behavior.," Journal of Applied Data Sciences., vol. 5, no. 1, pp. 100-113, 2024.

[11] Dinesh, Paidipati, A. S. Vickram, and P. Kalyanasundaram, "Medical image prediction for diagnosis of breast cancer disease comparing the machine learning algorithms: SVM, KNN, logistic regression, random forest and decision tree to measure accuracy," AIP Conference Proceedings, vol. 2853, no. 1, 2024.

[12] Mridula, B., A. Hency Juliet, and N. Legapriyadharshini, "Deciphering Social Media Sentiment for Enhanced Analytical Accuracy: Leveraging Random Forest, KNN, and Naive Bayes," 10th International Conference on Communication and Signal Processing (ICCSP), 2024.

[13] Ramdas, Soorya, and Neenu NT Agnes, "Leveraging Machine Learning for Fraudulent Social Media Profile Detection.," Cybernetics and Information Technologies., vol. 24, no. 1, pp. 118-136, 2024.

[14] Sudhakar, M and K. P. Kaliyamurthie, "Detection of fake news from social media using support Vector machine learning algorithms," Measurement: Sensors, vol. 32, p. 101028, 2024.

[15] Dixit, Dheeraj Kumar, Amit Bhagat, and Dharmendra Dangi,"A comparative analysis for detecting fake news using supervised learning algorithms," AIP Conference Proceedings, vol. 2900, no. 1, 2024.

[16] Choudhury, Arghya, Arpita Mondal, and Subhadeep Sarkar, "Searches for the BSM scenarios at the LHC using decision tree-based machine learning algorithms: a comparative study and review of random forest, AdaBoost, XGBoost, and LightGBM frameworks.," The European Physical Journal Special Topics, pp. 1-39, 2024.

[17] Ontor, Md Risalat Hossain, et al. "Leveraging Digital Transformation and Social Media Analytics for Optimizing US Fashion Brands' Performance:A Machine Learning Approach.," American Research Index Library, pp. 45-56, 2024.