**SENTIMENT ANALYSIS OF TWEETS DURIN KENYA'S 2022 PRESIDENTIAL ELECTIONS**

**Introduction**

Social media has become an integral part of modern-day political campaigns with political candidates and parties using platforms such as Twitter, Facebook, and Instagram to reach and engage with voters.

Social media can impact election outcomes in various ways such as through campaigning, voter outreach, public opinion analysis, misinformation spread, and targeted advertising. However, the impact of social media on elections is not yet fully understood.

This research will use machine learning algorithms to explore the relationship between social media (Twitter) activity and election outcomes, providing insights into the influence of social media (Twitter) on election campaigns.

**Problem Statement**

The widespread use of social media especially Twitter in recent years has dramatically changed the way political campaigns reach and engage with potential voters.

However, despite its widespread use, there is limited research on the extent to which Twitter influences election outcomes and the role of data science techniques in monitoring public opinion and misinformation on social media platforms.

The increasing use of social media in political campaigns has also raised concerns about the spread of misinformation, fake news, and biased algorithms.

The purpose of this study is to address this gap in the literature by examining the influence of social media on election campaigns and the role of data science techniques in monitoring public opinion and misinformation.

This study will help to better understand the sentiments shared on Twitter duringg Kenya Election 2022 . The findings of this study will have important implications for political campaigns, providing valuable insights into the role of social media and data science in modern political campaigns.

**Objectives of the Study**

The primary objective of this research is to study the relationship between social media (Twitter) activity and election outcomes.

The following specific objectives will be addressed:

1. To study twitter users based on their tweets to determine their political party preference
2. To segment twitter users into distinct groups based on their political party preference
3. To provide insights into the influence of social media on election campaigns and identify any patterns or trends that emerge.

**Experimental Design**

The research will employ a mixed-methods approach, including both qualitative and quantitative methods.

Data will be collected from Twitter social media platform using APIs or web scraping techniques.

The collected data will be preprocessed and analyzed using machine learning algorithms such as classification, clustering, and regression. The performance of the models will be evaluated using metrics such as accuracy, precision, recall, and F1 score.

**About Datasets**

We scraped the data from Twitter using APIs. 4070 data records were scraped from Twitter.

The following features were scraped:

tweet.user.verified,

tweet.user.created,

tweet.user.followersCount,

tweet.user.friendsCount,

tweet.retweetCount,

tweet.lang,

tweet.date,

tweet.likeCount,

tweet.sourceLabel,

tweet.id,

tweet.content,

tweet.hashtags,

tweet.conversationId,

tweet.inReplyToUser,

tweet.coordinates,

tweet.place,

tweet.retweetCount]

**Modelling**

We built the model using Logistic Regression and Support Vector Machine algorithms. Between the two, Logistic Regression performed better with 61.22%.