International Journal of
Technology and Systems
(IJTS)

Framework for Mitigating Phishing E-mail in the Kenyan Banking Industry Using Artificial Intelligence (AI)

Asiema Mwavali

Framework for Mitigating Phishing E-mail in the Kenyan Banking Industry Using Artificial Intelligence (AI)

Asiema Mwavali
Department of Computing and Informatics, Technical University of Kenya, Nairobi

Article History

Received 8th May 2024

Received in Revised Form 14th June 2024

Accepted 17th July 2024

How to cite in APA format:
Mwavali, A. (2024). Framework for Mitigating Phishing E-mail in the Kenyan Banking Industry Using Artificial Intelligence (AI). *International Journal of Technology and Systems*, *9*(2), 40–66. https://doi.org/10.47604/ijts.2781

Abstract

Purpose: Phishing is a significant cybercrime threat that affects individuals and organizations globally, including the banking industry in Kenya. The sophistication of phishing attacks continues to increase, and it is increasingly challenging traditional security measures to mitigate these threats. The purpose of this thesis is to build a framework for mitigating phishing e-mail attacks in the banking industry in Kenya using artificial intelligence. Phishing emails are among the most common techniques of cyber-attacks utilized by assailants to gain unauthorized access to sensitive information such as financial details, personal information, and login credentials. These attacks can have devastating effects on the victims, leading to financial loss, reputation damage, and even identity theft.

Methodology: The framework development consists of four main stages: data collection, data preprocessing, model training, and deployment. In the data collection stage, a dataset of phishing and non-phishing emails is gathered from various sources such as public databases, dark web forums, and bank employees mail. In the data preprocessing stage, the collected data is cleaned, preprocessed, and labeled. In the model training stage, machine learning algorithms and NLP techniques is used to develop a robust phishing and non-phishing emails detection model. In the deployment stage, the model is integrated into the bank's email system to detect and block phishing emails in real-time. The framework is then evaluated using a dataset of phishing and non-phishing e-mails collected from the banking industry in Kenya. Various metrics such as accuracy, precision, recall, and F1-score are used to evaluate the framework. The framework is able to detect new phishing e-mails that were not previously included in the dataset, demonstrating its ability to adapt to new threats.

Findings: The framework is based on a hybrid approach that combines machine learning algorithms, natural language processing (NLP) techniques, and human expertise that identify and prevent phishing emails from reaching their targets. The four main components of this framework include e-mail filtering, feature extraction, classification, and response. The e-mail filtering component uses several algorithms to identify and filter suspicious e-mails. The feature extraction component analyzes the content of the e-mail and extracts relevant features to help classify the e-mail as either legitimate or phishing. The classification component uses machine-learning algorithms to classify the e-mail as either legitimate or phishing. Finally, the response component takes appropriate action based on the classification results.

Unique Contribution to Theory, Practice and Policy: The framework provides an effective way to identify and mitigate phishing e-mail attacks, reducing the risk of data breaches and financial losses.

Keywords: *Phishing Email, Cyber Threats, Mitigation, Banking Industry, Machine Learning Algorithm, Natural Language Processing (NLP) Techniques, E-Mail Filtering, Feature Extraction, Classification, Accuracy, Precision, Recall, F1-Score*

INTRODUCTION

Phishing is the process of utilizing social engineering procedure by the attackers whose aim is to gather targeted victim confidential data and install mischievous program (malware) on their computer systems (Andress (2019). For this to be achieved, the targeted client is persuaded or tricked to click a mischievous link within email, which diverts the client into a spoofed website that is erected for sole intention of amassing subtle data, such as identifications. The impersonated websites utilized in phishing characteristically looks similar to those familiar with websites, social media, organization, or targeted bank industries individual's websites. However, part of these websites might appear clearly forgery with;

Challenges

Phishing emails are a major threat to Kenyan banking industry. In recent years, there has been a significant increase in the number and sophistication of phishing attacks. These attacks have resulted in significant financial losses for banks and their customers.

Existing phishing email mitigation frameworks and models have limitations that make them vulnerable to new and emerging phishing attacks. These limitations include:

- **Inability to detect and block new and emerging phishing attacks:** Phishing attacks are constantly developing new techniques to evade existing detection and prevention mechanisms.
- **High false positive rates:** Existing phishing mitigation framework and models often generate a high number of false positives. This can lead to legitimate emails being blocked or flagged as spam, which can be disruptive for users.
- **Lack of customization for the Kenyan banking industry:** Many existing phishing email mitigation frameworks and models are not specifically designed for Kenyan banking industry. This means that they may not be effective at detecting and blocking phishing attacks that are targeted at Kenyan banks and their customers.
- **PwC Kenya Economic Crime Survey 2023.** This survey provides insights into the economic crime risks that are facing Kenyan businesses. It found that email phishing is the most common type of economic crime, accounting for 48% of all reported incidents in 2022. https://www.pwc.com/ke/en/publications/economic-crime-survey.html
- **KPMG Kenya Fraud and Risk Survey 2022.** This survey provides insights into the fraud risks that are facing Kenyan businesses. It found out that email phishing is one of the most common types of fraud, accounting for 37% of all reported fraud incidents in 2021.

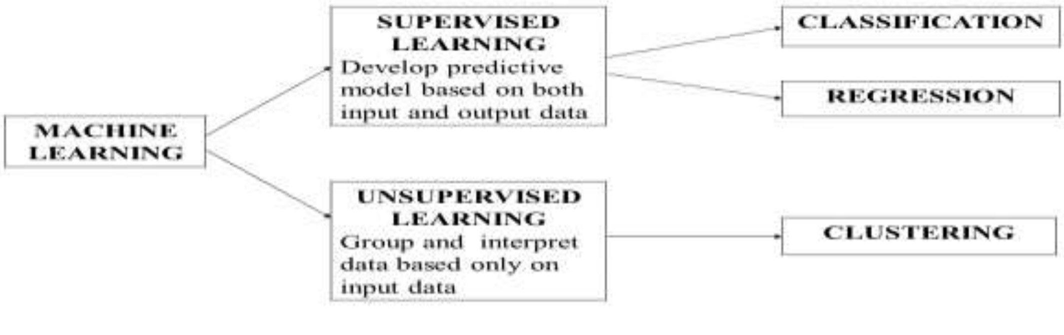
Artificial intelligence (AI) has the potential to address these limitations and develop more effective phishing email mitigation frameworks and models. AI can be used to develop models that can learn and adapt to new emerging phishing email attacks. AI can also be used to develop models that are more specific to the Kenyan banking industry.

The goal of this research project is to develop an AI-powered phishing email mitigation framework for Kenyan banks. This framework will address the limitations of existing frameworks and models, and it will be specifically designed to detect and block phishing attacks that are targeted at Kenyan banks and their customers.

Resolution

AI-based frameworks can provide real-time analysis of incoming emails, detect and adapt to changing attack patterns and reduce the number of false positives.

By using AI to analyze large amounts of data (Big Data) and identify patterns and anomalies, banks can effectively detect and block phishing emails, which can help to reduce the risk of successful phishing attacks and protect their customers' information. A combination of the Supervised Learning and Unsupervised Learning approaches can be used to maximize effectiveness of mitigating phishing email threats.


MACHINE LEARNING
SUPERVISED LEARNING
Develop predictive model based on both input and output data
UNSUPERVISED LEARNING
Group and interpret data based only on input data
CLASSIFICATION
REGRESSION
CLUSTERING

Supervised Learning verses Unsupervised Learning (Mathworks, n.d.)

Figure 1: Supervised Learning verses Unsupervised Learning

Unsupervised Learning

This research is going to adopt Unsupervised Learning approaches where it involves training AI model using unlabeled data. The model identifies the structures, patterns and characteristics in the data on its own. Unsupervised Learning can identify previous unseen phishing attacks by clustering similar patterns in the data. It groups the emails that share similarities and characteristics, which could be further, analyzed. The source of data for this research will be emails. And clustering technique or algorithm will be used, as this will involve grouping of data points into similar or related clusters based on similarities.

Clustering Algorithm

Clustering Algorithm will be used for this task. Clustering is a machine-learning technique that involves grouping similar data points into clusters based on their similarities or differences. The goal of the clustering is to identify patterns and grouping based on their features (like the links, subject line, any attachments, and the body of the email) of which may not be immediately apparent to human observers.

Clustering

Apply a clustering algorithm to the preprocessed dataset of emails. One popular algorithm that can be used for this purpose is the k-means algorithm. The number of clusters can be determined by using techniques such as the elbow method or silhouette score.

Cluster Analysis

Analyze the resulting clusters to identify similarities and differences between the emails. This can be done by visualizing the clusters using techniques such as t-SNE or PCA. The clusters can also be manually inspected to identify any patterns or characteristics that are indicative of phishing attacks.

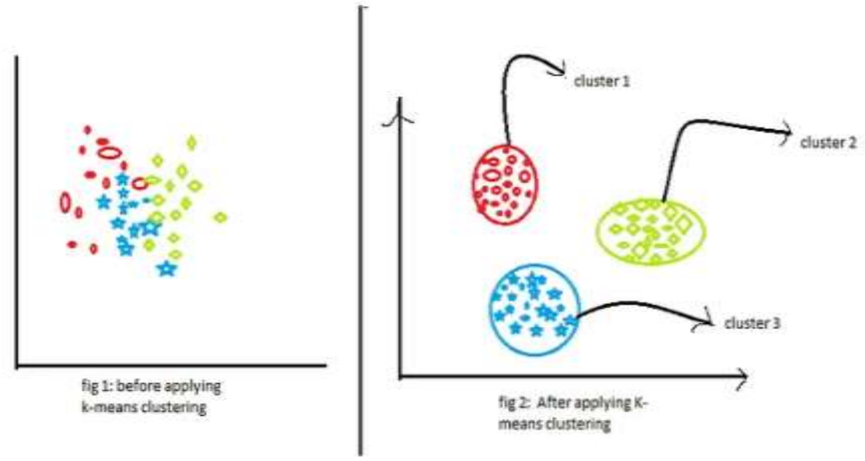


Figure 2: Illustration Showing Unsupervised Machine Learning Techniques

How Artificial Intelligence (AI) can be used to Mitigate Phishing of E-Mail

- Detection of phishing attacks: AI can be used to detect and identify phishing attacks in real-time.
- Email Filters: AI algorithms can be used to create email filters that can detect phishing emails and flag them as spam or potentially harmful.
- Behavioral Analysis: AI can be used to analyze the behavior of users and identify anomalies in their behavior that may indicate a phishing attempt.
- Predictive Analytics: AI can be used to analyze data from past phishing attempts to predict the likelihood of future attacks and identify areas where security can be improved.
- Training and awareness: AI can also be used to train employees on how to recognize and respond to phishing attacks.
- Natural Language Processing (NLP): AI can be used to analyze the text content of emails and identify suspicious patterns of language or grammar that may indicate a phishing attempt.

Threats and Vulnerabilities in the Financial Sector

Sacco cybersecurity (2018) accounts indicate that data breaches like abuse of privileged access, critical data manipulation, phishing email attacks and insider assaults are most common assaults pursuing organizations in the financial sectors (Serianu Limited, 2018). There have been numerous cases of data breaches within the financial sectors globally. For instance, in 2014 there was a report of phishing email bout on JP Morgan Chase and Co. occasioned in exfiltration of data from households estimated at 76 million and SMEs projected at 7 million. There were no accounts of any monetary loss; this was confirmed by bank administrators though there was concern that stolen data could be used to unveiling assaults in future on the customers who were affected. (Henley, 2019).

LITERATURE REVIEW

Recent Cases of Phishing Attacks

Kenyan E-Citizen Platform Hack (2023)

In 2023, the e-Citizen platform, a vital online portal for accessing government services in Kenya, was compromised by a sophisticated cyberattack. The attack involved a combination of phishing emails and social engineering techniques to gain access to user credentials and sensitive personal information. This incident highlights the vulnerability of critical infrastructure and the need for robust security measures to protect sensitive data. (BBC July 2023)

Artificial Intelligence (AI)

Machine Learning (ML) and Artificial Intelligence (AI) are closely correlated concepts, but then they are not similar. Machine learning is a specific subfield of AI that focuses on building systems that can learn from data and improve their performance over time.

AI on the other hand, refers to the overall field of computer science and engineering that aims to create machines and systems that can perform tasks that would normally require human intelligence, such as recognizing speech, making decisions, and identifying patterns.

AI/ML models bid flexibility likened to traditional statistical and econometric frameworks, can aid search otherwise hard-to-detect relationships between variables, and amplify the toolkits used by institutions. Evidence suggests that AI methods often outperform linear regression-based methods in forecast accuracy and robustness (Bolhuis and Rayner 2020).

Relating Artificial Intelligence (AI) to Cyber Security Awareness

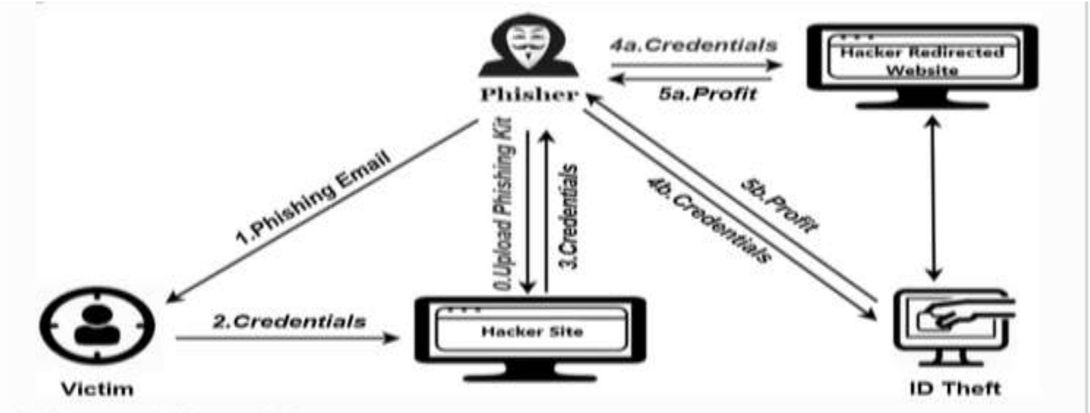
Artificial intelligence (AI) and cyber security awareness are related in numerous ways:

1. AI-based security solutions: AI may be used to examine vast volumes of data and spot trends and abnormalities that can be used to recognize and thwart online threats including phishing emails, malware, and network intrusions.
2. AI can be applied to the development of dynamic and interesting cyber-security awareness training courses.
3. Real-time monitoring and incident response: AI can monitor the network and email system in real-time to spot and notify users of any suspicious activity.
4. Increasing user awareness: AI can be used to offer immediate assistance to users who have been the target of phishing attempts.
5. Improving security compliance: AI can be used to monitor and enforce security compliance, including identifying regulatory standards that are not being followed and flagging suspicious activity.

Problem with the Current Framework

The framework's effectiveness and widespread acceptance will be greatly influenced by how the following issues with the current body of knowledge are resolved. When building an AI-based framework for phishing email mitigation, there are various restrictions and difficulties that must be taken into account and these include the following:

- Big Data / Data accessibility: The necessitates of having access to a significant amount of recent, high-quality training data might be challenging to find.
- Integration with Existing Systems: Integrating the framework with already-existing systems, like email servers, can be quite difficult.
- Technological / User Acceptance: The framework must be simple to use, user-friendly, and satisfying to use.
- Attitude / False Positives and Negatives: Avoiding false positives and false negatives, which can have detrimental effects, is a big problem.



Phishing attack diagram [26]

Figure: Conceptual Framework

METHODOLOGY

Introduction

The research methodology for this project will involve the following steps:

Literature review: A comprehensive review of the existing literature on phishing email attacks and AI-based phishing email mitigation frameworks was conducted to identify the state-of-the-art and the research gaps. In particular, the literature review focused on studies that have been conducted in the Kenya. For example, one relevant study is the "Kenya Banks Association (KBA) Cyber Security Report 2022," which found that phishing emails are one of the most common types of cyber threats facing Kenyan banks. Another relevant study is the "Central Bank of Kenya (CBK) Fraud Report 2021," which found that phishing emails accounted for 35% of all reported fraud incidents in the Kenyan banking sector in 2020.

Area of Study

The use of unsupervised machine learning techniques to reduce phishing emails in the banking industry in Kenya was the subject of this study. The study concentrated on using AI to categorize and recognize phishing emails automatically based on their traits and trends.

Data gathering, data preprocessing, method choice, and evaluation of the suggested fix were all parts of the research process.

Paradigm of Research

The paradigm of research for this project is **design science research (DSR)**. DSR is a research paradigm that focuses on the design and evaluation of artifacts to solve real-world problems. In this project, the artifact is an AI-powered phishing email mitigation framework for Kenyan banks.

The DSR paradigm was well-suited for this project because it allowed for the development and evaluation of the AI-powered framework in a systematic and rigorous manner. The DSR process typically involved the following steps:

- **Identified the problem:** The first step was to identify the problem that required to be solved. In this case, the problem was the increasing number and sophistication of phishing email attacks targeting Kenyan banks.
- **Designed the artifact:** The next step was to design the artifact that solved the problem. In this case, the artifact was the AI-powered phishing email mitigation framework.
- **Evaluated the artifact:** The third step was to evaluate the artifact to ensure that it met the requirements and solved the problem. In this case, the AI-powered framework was evaluated using publicly available datasets of phishing emails and surveys and interviews with Kenyan bank customers and cybersecurity experts.
- **Deployed the artifact:** Once the artifact has been evaluated and found to be effective, it can be deployed in the real world. In this case, the AI-powered framework could be deployed at Kenyan banks to help mitigate phishing email attacks.

Research Design

A quantitative research approach was used in this study, specifically an experimental one. An experimental design was appropriate for this investigation because it enabled variable manipulation and the definition of cause-and-effect linkages. The framework for phishing email mitigation in the banking business utilizing AI was the independent variable in this study, while the success of the framework in phishing attack mitigation was the dependent variable.

There were two phases to the investigation. The suggested framework was created utilizing unsupervised learning methods in the initial stage. Using a dataset of phishing emails, the framework's performance was assessed in the second phase by contrasting its rate of phishing detection with that of other solutions. Performance indicators including precision, recall, and F1 score was used in the study to assess the usefulness of the suggested framework. These metrics was used to assess how well the suggested framework performs in comparison to current approaches. In order to depict the categorization outcomes achieved by the framework, the study also used a confusion matrix.

Moreover, a control group and an experimental group was included in the research design. A sample of phishing emails chosen at random and categorized using existing tools made up the control group. The same sample of phishing emails that were used to classify the control group made up the experimental group. To assess the efficacy of the suggested framework, the performances of the two groups were contrasted.

Data Collection

Data collection was collected from a variety of sources that included:

- Publicly available datasets of phishing emails, such as the Phishing.org Phishing Test Email Dataset
- Phishing emails collected from Kenyan banks, with the permission of the banks
- Surveys and interviews with Kenyan bank customers and cybersecurity experts

The surveys and interviews were designed to collect data on the following:

I. The types of phishing emails that Kenyan bank customers were most likely encountered

II. The methods that Kenyan banks were using to mitigate phishing email attacks

III. The challenges that Kenyan banks were facing in mitigating phishing email attacks

IV. The expectations of Kenyan bank customers for AI-based phishing email mitigation frameworks

Data preprocessing and analysis: The collected data was preprocessed and analyzed using a variety of machine learning and artificial intelligence techniques to identify the key features and patterns associated with phishing emails. The preprocessing and analysis were tailored to the specific characteristics of phishing emails in the Kenya. For example, the preprocessing involved removing Kenyan-specific slang and colloquialisms from the phishing emails, and the analysis involved identifying Kenyan-specific phishing attack techniques.

Data Collection Tools and Methods

To collect qualitative data from Kenyan bank customers and cybersecurity experts, the following tools and methods were used:

Kenyan Bank Customers

Surveys: an online survey was developed to collect data on customers' experiences with phishing emails, their perceptions of the risks of phishing attacks, and their expectations for AI-based phishing email mitigation frameworks. The survey distributed to a sample of Kenyan bank customers using a variety of methods, such as email, social media, and in-bank recruitment.

Interviews: in-depth interviews were conducted with a small group of Kenyan bank customers to get more detailed feedback on their experiences with phishing emails and their thoughts on the proposed AI-powered framework. The interviews were conducted in person or online, depending on the preference of the participant.

Kenyan cybersecurity experts:

Interviews: I conducted interviews with cybersecurity experts at Kenyan banks to learn about the challenges they face in mitigating phishing email attacks and their feedback on the proposed AI powered framework. The interviews were conducted in person or online, depending on the preference of the participant.

To ensure the quality of my data collection process, I took the following steps:

- **Pilot testing surveys and interview guides:** I piloted test my survey and interview guides with a small group of people to ensure that they are clear, concise, and easy to understand.

- **Obtaining informed consent:** I obtained informed consent from all participants before collecting any data. This means that I provided them with information about the study, including its purpose, risks, and benefits, and gave them the opportunity to opt out.
- **Recording and transcribing interviews:** I recorded and transcribed all interviews so that I could accurately capture the participants' responses.
- **Analyzing the data using a qualitative data analysis software package:** I used a qualitative data analysis software package, such as FreeQDA or Coding Analysis Toolkit to analyze the data I collected. This software helped me to identify patterns and themes in the data and developed a deeper understanding of the participants' perspectives.

Sampling Size

Given the large population of bank customers in Kenya (approximately 64 million), conducting a study involving all customers was impractical and resource-intensive. Therefore, employing an appropriate sampling technique was crucial to ensure the representatives and generalization of the research findings.

Several sampling techniques was considered for this research project including **S**imple Random Sampling, Stratified Sampling, Cluster Sampling and Multistage Sampling.

Recommended Sampling Technique For this research

Considering the research objectives and the characteristics of the population, using **stratified sampling** was recommended for this project. This method allowed for a more representative sample by capturing the diversity of the customer base in terms of age, gender, and bank type.

The following steps outlined the recommended sampling procedure:

Defined Strata: Identified relevant characteristics that divided the population into strata. In this case, age, gender, and bank type was considered to be suitable strata.

Allocated Sample: Sample size was allocated proportionally to each stratum based on the relative size of each stratum within the population.

Selected Sample within Strata: Within each stratum, employed a simple random sampling method that selected the specified number of customers.

Determined Sample Size: Calculated appropriate sample size based on the research objectives and the desired level of precision. Yamane's formula was used for this purpose.

The number of customers in Kenyan banks was estimated to be over 64 million as of year 2023. This figure was based on data from the Central Bank of Kenya (CBK), which showed that there were approximately 66.3 million deposit accounts in commercial banks in Kenya in the year 2022.

Central Bank of Kenya. Bank Supervision & Banking Sector Reports. Retrieved November 14, 2023 from https://www.centralbank.go.ke/reports/bank-supervision-and-banking-sector-reports/

CBK does not track the number of individual bank customers. However, it does provide data on the number of deposit accounts, which can be used to approximate the number of customers.

n = N / (1 + N * e^2)

n = 64 million / (1 + 64 million * 0.05^2) n = 400

Sample Size Calculation

The sample size for this research project was calculated using the following **Yamane's formula:**

n = N / (1 + N * e^2)

Where:

n is the sample size

N is the population

Size **e** is the margin of **error**

Given that the population size of Kenyan bank customers is **64** million and I want a margin of **error** of **5%**, the sample size would be calculated as follows:

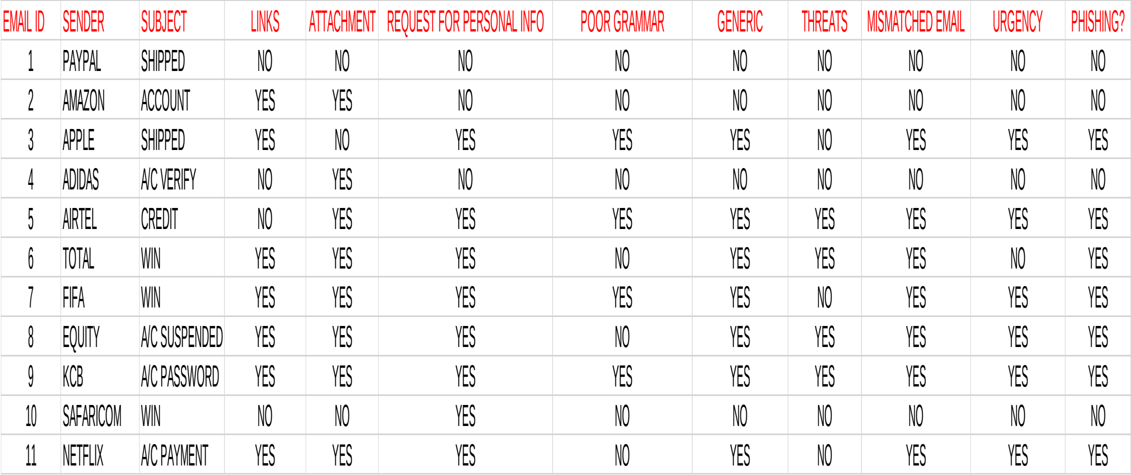
n = 64 million / (1 + 64 million *0.05^2) = 400

This means that I needed to collect data from a sample of at least 400 Kenyan bank customers in order to achieve the desired level of accuracy and reliability.

Extraction of Feature

Extraction of Feature is a vital step in any machine learning project. In this study, feature extraction was done on the dataset to identify the most relevant features that could be used in the unsupervised learning model for detecting phishing emails. The aim of extraction of feature was to lessen the dimensionality of dataset by selecting the most informative features. Extraction of feature approaches such as Independent Component Analysis (ICA), Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA), was explored in this study.

Table 1: A sample Illustrating Extraction of Feature, Phishing and Non-Phishing Emails



EMAIL ID	SENDER	SUBJECT	LINKS	ATTACHMENT	REQUEST FOR PERSONAL INFO	POOR GRAMMAR	GENERIC	THREATS	MISMATCHED EMAIL	URGENCY	PHISHING?
1	PAYPAL	SHIPPED	NO	NO	NO	NO	NO	NO	NO	NO	NO
2	AMAZON	ACCOUNT	YES	YES	NO	NO	NO	NO	NO	NO	NO
3	APPLE	SHIPPED	YES	NO	YES	YES	YES	NO	YES	YES	YES
4	ADIDAS	A/C VERIFY	NO	YES	NO	NO	NO	NO	NO	NO	NO
5	AIRTEL	CREDIT	NO	YES	YES	YES	YES	YES	YES	YES	YES
6	TOTAL	WIN	YES	YES	YES	NO	YES	YES	YES	NO	YES
7	FIFA	WIN	YES	YES	YES	YES	YES	NO	YES	YES	YES
8	EQUITY	A/C SUSPENDED	YES	YES	YES	NO	YES	YES	YES	YES	YES
9	KCB	A/C PASSWORD	YES	YES	YES	YES	YES	YES	YES	YES	YES
10	SAFARICOM	WIN	NO	NO	YES	NO	NO	NO	NO	NO	NO
11	NETFLIX	A/C PAYMENT	YES	YES	YES	NO	YES	NO	YES	YES	YES

Model Choice

In this study, phishing emails in the banking industry was reduced using an unsupervised learning methodology. Clustering algorithms were a good choice among unsupervised learning methods for this use. In this project, I explored K-Means. K-Means is a well-known unsupervised learning method that is commonly used for clustering problems. By reducing the distance between data points and their assigned cluster centroids, the method divided the data into K clusters. The algorithm allocated data points to their nearest centroid iteratively and updated the centroid until the clusters stopped changing or the maximum number of iterations was reached.

The reasoning behind employing K-Means for this phishing email detection project was that it efficiently grouped similar sorts of emails together basing on their characteristics. This allowed for the identification of trends and characteristics specific to phishing emails, which was then used to construct successful phishing email detection models.

Furthermore, it was demonstrated to perform well on a number of datasets and had been utilized in a variety of applications such as picture segmentation and recommendation systems. S. Biswas and R. Panigrahi (2021)

Model Training

The specifics of the model training procedure were covered in this section. Unsupervised learning was a component of the suggested method, therefore clustering algorithms like k-means or hierarchical clustering was used. The previously gathered data was preprocessed and converted into a format that the clustering algorithms could use. By comparing various clustering models and choosing the one that yielded the best results, the number of clusters was decided.

Model Development and Evaluation

AI-based phishing email classification models was developed and evaluated using the preprocessed and analyzed data. The models were trained on a subset of the data and evaluated on a held-out test set. The evaluation was conducted using a variety of metrics, including accuracy, precision, recall, and F1 score.

Framework development and evaluation: An AI-powered phishing email mitigation framework was developed using the developed AI models. The framework was designed to be specific to the needs of Kenyan banks and their customers. For example, the framework was integrated with existing phishing email mitigation systems that were used by Kenyan banks. The framework was evaluated using a variety of metrics, including:

- The ability to detect and block phishing emails that are targeted at Kenyan banks and their customers.
- The ability to reduce the number of false positives.
- The ease of use.

DATA ANALYSIS AND RESULTS

Data Preprocessing and Feature Engineering.

Data preprocessing and feature engineering are crucial steps in preparing the data for analysis and training the AI model effectively. These steps involve transforming the raw data into a format suitable for the chosen algorithm and extracting relevant features that help the model distinguish between phishing and legitimate emails.

Data Cleaning and Transformation

Data cleaning involves identifying and correcting errors or inconsistencies within the data. This may include:

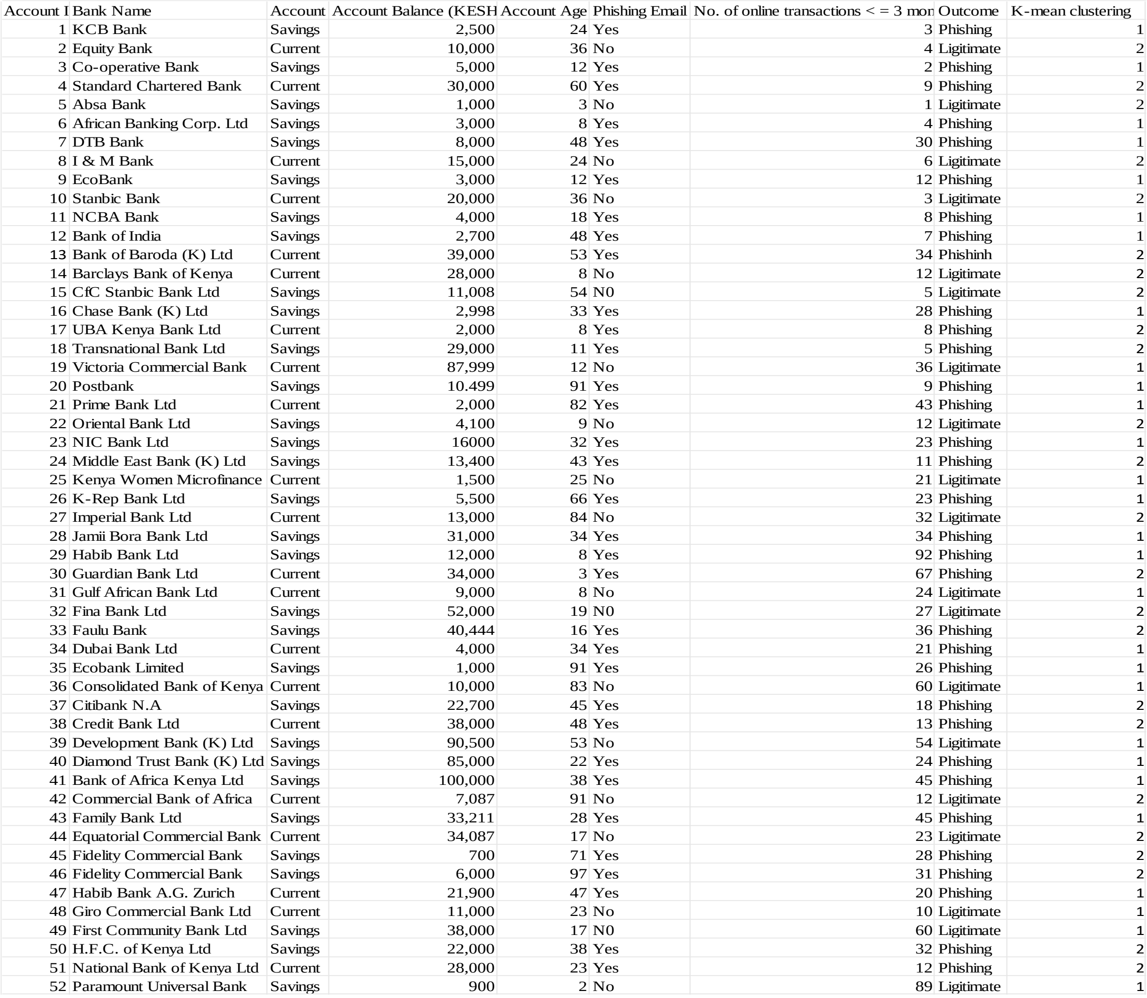
Removing duplicates: Eliminating redundant email entries to improve the overall data quality.

Handling missing values: Identifying and imputing missing values based on statistical methods or domain knowledge.

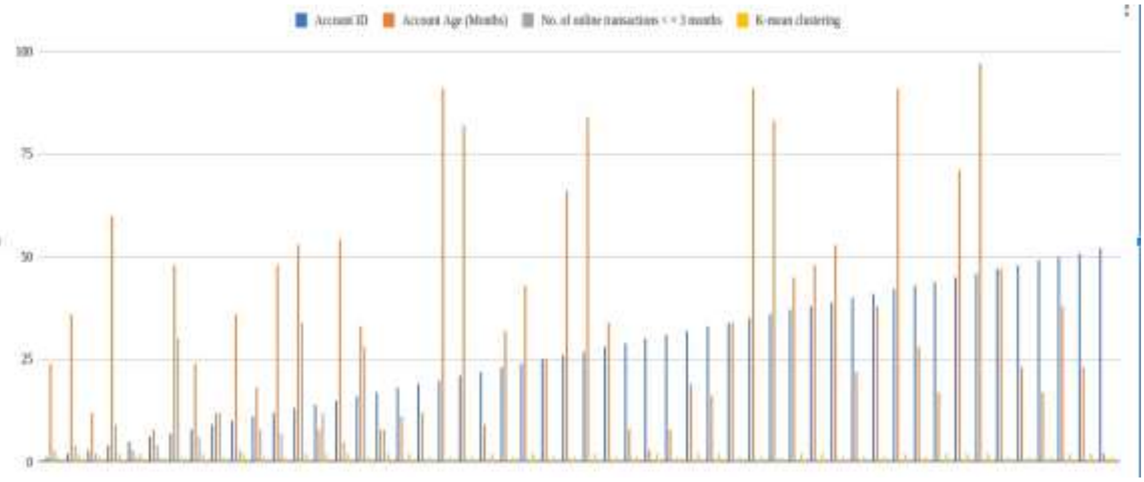
Formatting text data: Converting text to lowercase, removing punctuation and special characters, and normalizing words to their base form (stemming or lemmatization).

Standardizing numerical data: Scaling numerical features to a common range (e.g., z-score normalization) to ensure equal weightage during analysis.

Feature Selection and Extraction



Account I	Bank Name	Account	Account Balance (KESH	Account Age	Phishing Email	No. of online transactions < = 3 mon	Outcome	K-mean clustering
1	KCB Bank	Savings	2,500	24	Yes	3	Phishing	1
2	Equity Bank	Current	10,000	36	No	4	Ligitimate	2
3	Co-operative Bank	Savings	5,000	12	Yes	2	Phishing	1
4	Standard Chartered Bank	Current	30,000	60	Yes	9	Phishing	2
5	Absa Bank	Savings	1,000	3	No	1	Ligitimate	2
6	African Banking Corp. Ltd	Savings	3,000	8	Yes	4	Phishing	1
7	DTB Bank	Savings	8,000	48	Yes	30	Phishing	1
8	I & M Bank	Current	15,000	24	No	6	Ligitimate	2
9	EcoBank	Savings	3,000	12	Yes	12	Phishing	1
10	Stanbic Bank	Current	20,000	36	No	3	Ligitimate	2
11	NCBA Bank	Savings	4,000	18	Yes	8	Phishing	1
12	Bank of India	Savings	2,700	48	Yes	7	Phishing	1
13	Bank of Baroda (K) Ltd	Current	39,000	53	Yes	34	Phishinh	2
14	Barclays Bank of Kenya	Current	28,000	8	No	12	Ligitimate	2
15	CfC Stanbic Bank Ltd	Savings	11,008	54	N0	5	Ligitimate	2
16	Chase Bank (K) Ltd	Savings	2,998	33	Yes	28	Phishing	1
17	UBA Kenya Bank Ltd	Current	2,000	8	Yes	8	Phishing	2
18	Transnational Bank Ltd	Savings	29,000	11	Yes	5	Phishing	2
19	Victoria Commercial Bank	Current	87,999	12	No	36	Ligitimate	1
20	Postbank	Savings	10.499	91	Yes	9	Phishing	1
21	Prime Bank Ltd	Current	2,000	82	Yes	43	Phishing	1
22	Oriental Bank Ltd	Savings	4,100	9	No	12	Ligitimate	2
23	NIC Bank Ltd	Savings	16000	32	Yes	23	Phishing	1
24	Middle East Bank (K) Ltd	Savings	13,400	43	Yes	11	Phishing	2
25	Kenya Women Microfinance	Current	1,500	25	No	21	Ligitimate	1
26	K-Rep Bank Ltd	Savings	5,500	66	Yes	23	Phishing	1
27	Imperial Bank Ltd	Current	13,000	84	No	32	Ligitimate	2
28	Jamii Bora Bank Ltd	Savings	31,000	34	Yes	34	Phishing	1
29	Habib Bank Ltd	Savings	12,000	8	Yes	92	Phishing	1
30	Guardian Bank Ltd	Current	34,000	3	Yes	67	Phishing	2
31	Gulf African Bank Ltd	Current	9,000	8	No	24	Ligitimate	1
32	Fina Bank Ltd	Savings	52,000	19	N0	27	Ligitimate	2
33	Faulu Bank	Savings	40,444	16	Yes	36	Phishing	2
34	Dubai Bank Ltd	Current	4,000	34	Yes	21	Phishing	1
35	Ecobank Limited	Savings	1,000	91	Yes	26	Phishing	1
36	Consolidated Bank of Kenya	Current	10,000	83	No	60	Ligitimate	1
37	Citibank N.A	Savings	22,700	45	Yes	18	Phishing	2
38	Credit Bank Ltd	Current	38,000	48	Yes	13	Phishing	2
39	Development Bank (K) Ltd	Savings	90,500	53	No	54	Ligitimate	1
40	Diamond Trust Bank (K) Ltd	Savings	85,000	22	Yes	24	Phishing	1
41	Bank of Africa Kenya Ltd	Savings	100,000	38	Yes	45	Phishing	1
42	Commercial Bank of Africa	Current	7,087	91	No	12	Ligitimate	2
43	Family Bank Ltd	Savings	33,211	28	Yes	45	Phishing	1
44	Equatorial Commercial Bank	Current	34,087	17	No	23	Ligitimate	2
45	Fidelity Commercial Bank	Savings	700	71	Yes	28	Phishing	2
46	Fidelity Commercial Bank	Savings	6,000	97	Yes	31	Phishing	2
47	Habib Bank A.G. Zurich	Current	21,900	47	Yes	20	Phishing	1
48	Giro Commercial Bank Ltd	Current	11,000	23	No	10	Ligitimate	1
49	First Community Bank Ltd	Savings	38,000	17	N0	60	Ligitimate	1
50	H.F.C. of Kenya Ltd	Savings	22,000	38	Yes	32	Phishing	2
51	National Bank of Kenya Ltd	Current	28,000	23	Yes	12	Phishing	2
52	Paramount Universal Bank	Savings	900	2	No	89	Ligitimate	1



Feature selection involves identifying the most relevant features from the data that contribute to the prediction of phishing emails. This helps to improve the model's performance and reduce computational burden. Feature selection techniques include:

Filter methods: These methods select features based on statistical measures such as correlation or information gain.

Wrapper methods: These methods use the AI model itself to evaluate the importance of features and select the most relevant ones.

Embedded methods: These methods integrate feature selection into the model training process, allowing for simultaneous feature selection and model learning.

Feature extraction involves creating new features from the existing data that capture more information about the emails. This can be done using techniques such as:

N-grams: Extracting sequences of n words to capture contextual information in the email body.

Bag-of-words: Representing the email body as a vector of word frequencies.

TF-IDF: Representing the email body as a vector of word frequencies weighted by their importance in the entire dataset and their individual documents.

Part-of-speech tagging: Identifying the grammatical role of each word in the email body.

Dimensionality Reduction

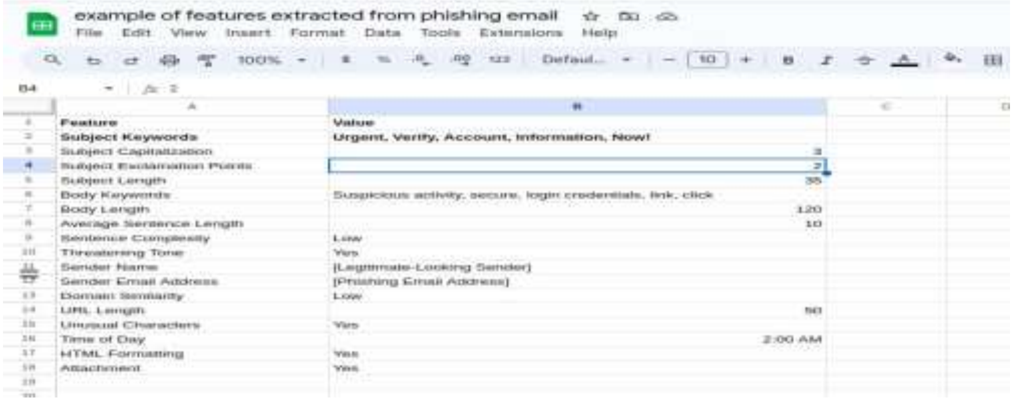
Dimensionality reduction techniques can be applied to reduce the number of features, especially when dealing with a large number of features. This can improve the efficiency of the algorithm and potentially reduce the risk of overfitting. Common dimensionality reduction techniques include:

Principal component analysis (PCA): This technique identifies the most important directions of variance in the data and projects the data onto these directions.

Linear discriminant analysis (LDA): This technique maximizes the separation between different classes (phishing and legitimate) while minimizing the dimensionality of the data.

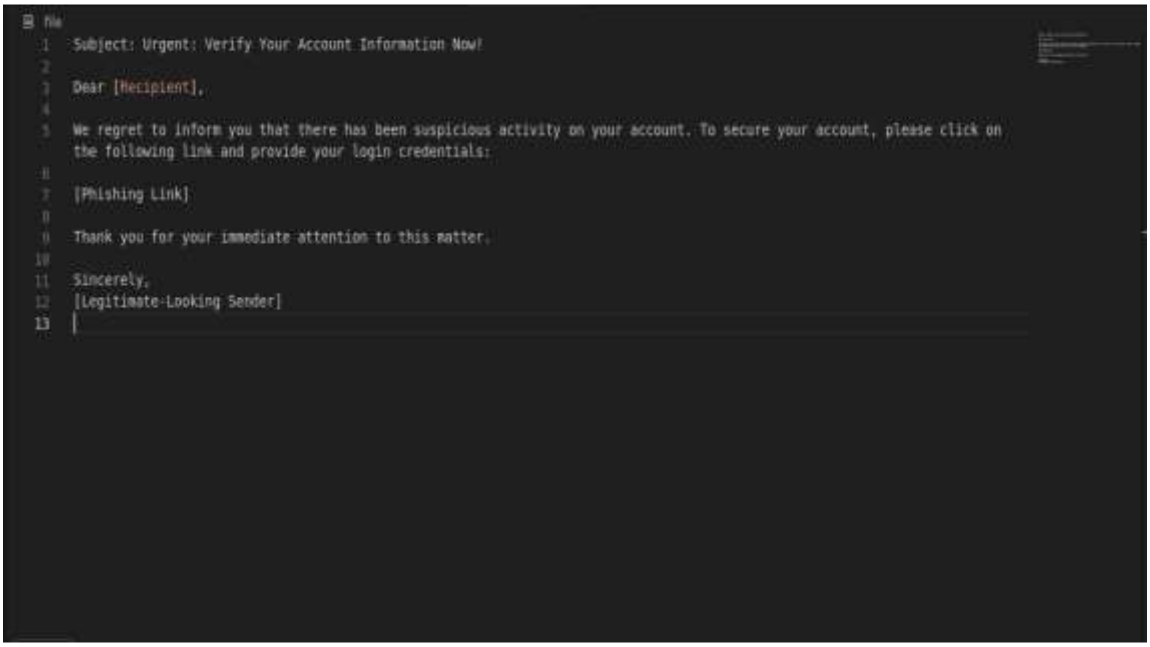
Example: Feature Extraction for Phishing Email Detection:

Original email.



example of features extracted from phishing email
File Edit View Insert Format Data Tools Extensions Help
100%
Defaul...
10
B4

Feature	Value
Subject Keywords	Urgent, Verify, Account, Information, Now!
Subject Capitalization	3
Subject Exclamation Points	2
Subject Length	35
Body Keywords	Suspicious activity, secure, login credentials, link, click
Body Length	120
Average Sentence Length	10
Sentence Complexity	Low
Threatening Tone	Yes
Sender Name	[Legitimate-Looking Sender]
Sender Email Address	[Phishing Email Address]
Domain Similarity	Low
URL Length	50
Unusual Characters	Yes
Time of Day	2:00 AM
HTML Formatting	Yes
Attachment	Yes

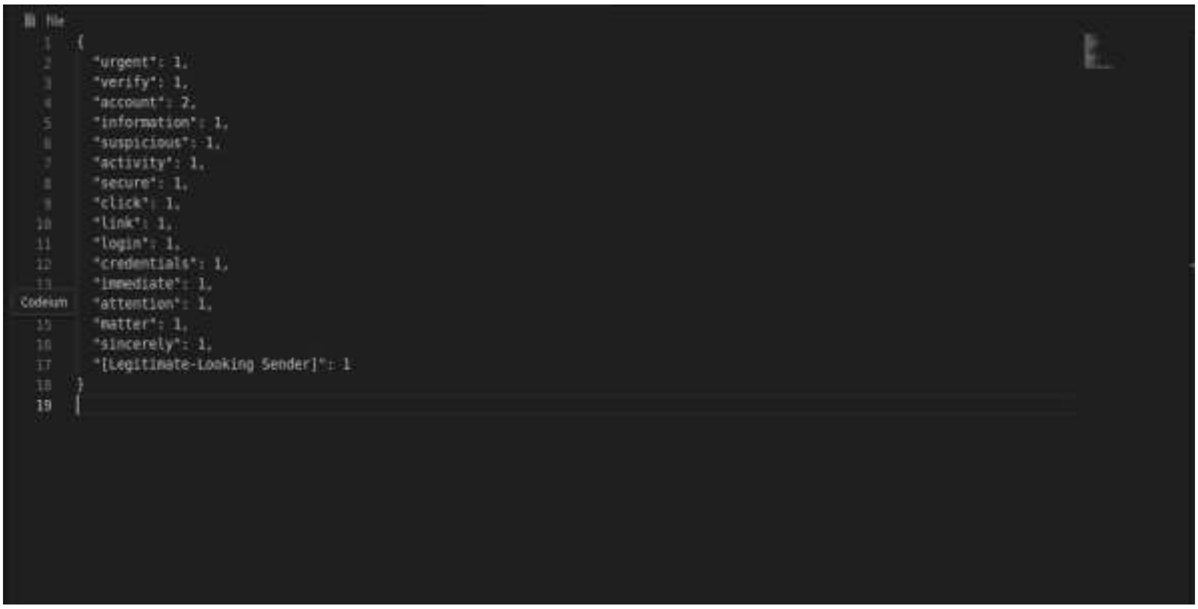


```
Subject: Urgent: Verify Your Account Information Now!

Dear [Recipient],

We regret to inform you that there has been suspicious activity on your account. To secure your account, please click on
the following link and provide your login credentials:

[Phishing Link]

Thank you for your immediate attention to this matter.

Sincerely,
[Legitimate-Looking Sender]
```

Feature Extraction Process

1. Word Distributions:

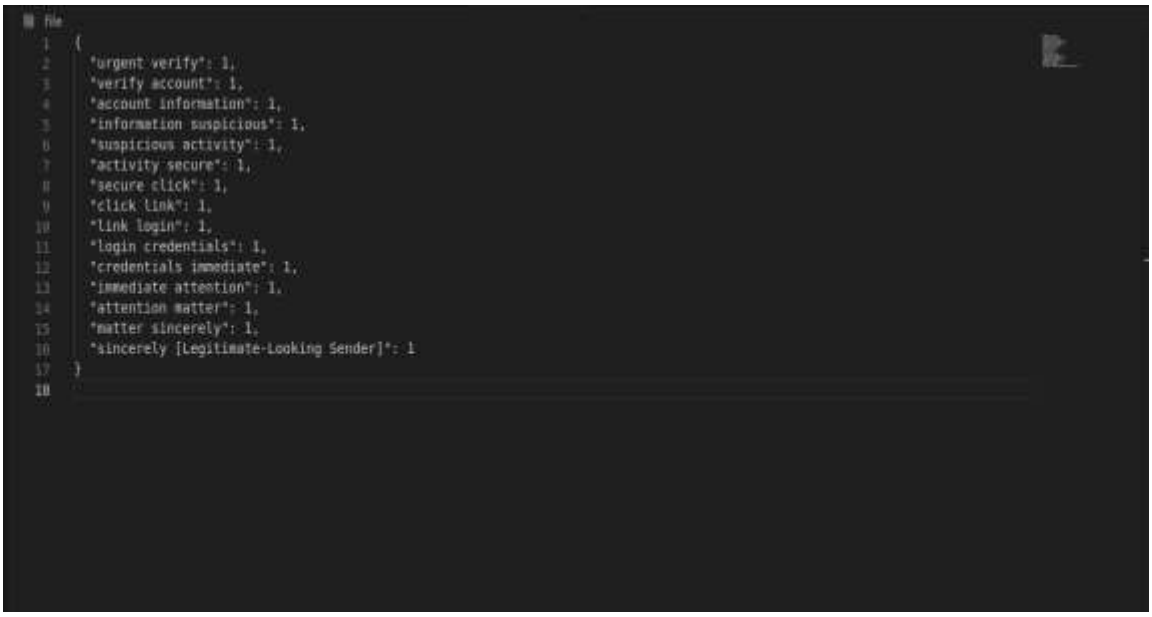
Determine how often each word appears in the email content. This records the word distribution and identifies possible keywords associated with phishing.



```
{
  "urgent": 1,
  "verify": 1,
  "account": 2,
  "information": 1,
  "suspicious": 1,
  "activity": 1,
  "secure": 1,
  "click": 1,
  "link": 1,
  "login": 1,
  "credentials": 1,
  "immediate": 1,
  "attention": 1,
  "matter": 1,
  "sincerely": 1,
  "[Legitimate-Looking Sender]": 1
}
```

2. N-grams

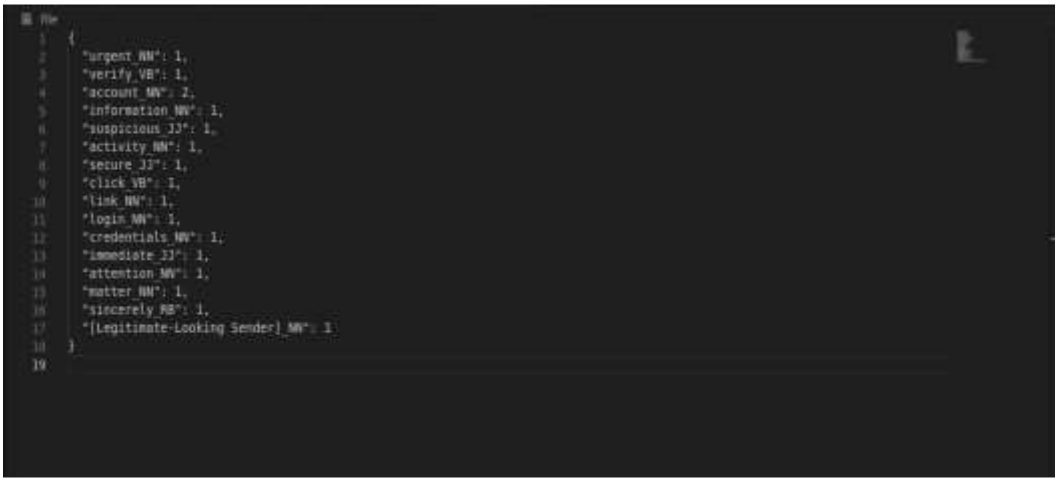
Take N consecutive word sequences and extract them. This facilitates the extraction of patterns and contextual data from the email text.



```
{
  "urgent verify": 1,
  "verify account": 1,
  "account information": 1,
  "information suspicious": 1,
  "suspicious activity": 1,
  "activity secure": 1,
  "secure click": 1,
  "click link": 1,
  "link login": 1,
  "login credentials": 1,
  "credentials immediate": 1,
  "immediate attention": 1,
  "attention matter": 1,
  "matter sincerely": 1,
  "sincerely [Legitimate-Looking Sender]": 1
}
```

3. Part-of-Speech Tags

Tag each word with its part of speech (e.g., noun, verb, adjective). This provides information about the grammatical structure and may reveal patterns associated with phishing.



```
{
  "urgent_NN": 1,
  "verify_VB": 1,
  "account_NN": 2,
  "information_NN": 1,
  "suspicious_JJ": 1,
  "activity_NN": 1,
  "secure_JJ": 1,
  "click_VB": 1,
  "link_NN": 1,
  "login_NN": 1,
  "credentials_NN": 1,
  "immediate_JJ": 1,
  "attention_NN": 1,
  "matter_NN": 1,
  "sincerely_RB": 1,
  "[Legitimate-Looking Sender]_NN": 1
}
```

Figures 1-3 illustrate an example of feature extraction for phishing email detection. The original email text is transformed into various features such as word frequencies, n-grams, and part-of-speech tags. These features are then used to train the AI model to distinguish between phishing and legitimate emails.

Training the AI Model

Once these features are extracted from a labeled dataset containing both phishing and legitimate emails, they can be used to train an AI model. The model learns to distinguish between the patterns associated with phishing and legitimate emails based on the feature representations.

This feature extraction process allows the model to capture both the content and context of the email, enabling it to generalize well to new, previously unseen phishing threats.

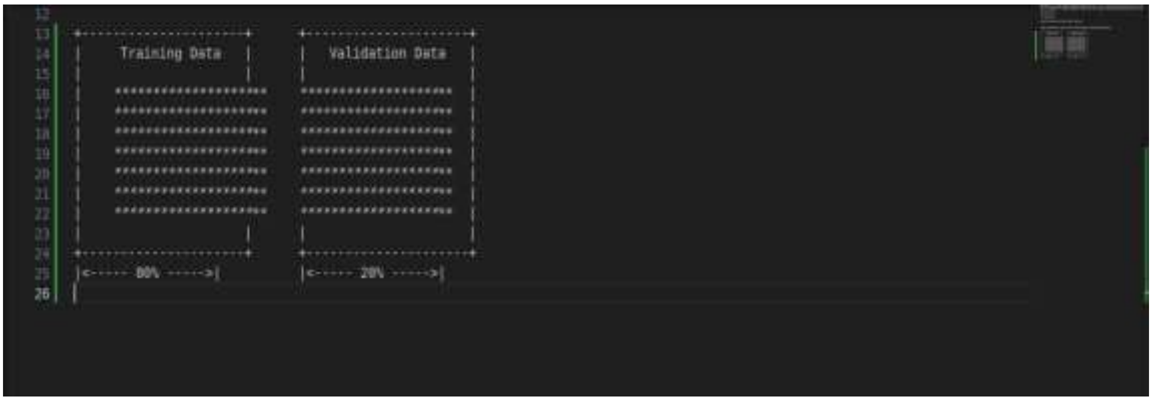
Model Training and Evaluation

Description of the AI Model

In this research project, we have adopted the use of an unsupervised learning and clustering algorithm, specifically K-means clustering, for phishing email detection. This algorithm groups emails based on their similarity in features without needing pre-labelled data, making it suitable for our scenario where we may not have a large labelled dataset of phishing and legitimate emails.

Training and Validation Data Split

Prior to training our K-means clustering model, we need to split our dataset into two separate subsets: training data and validation data. The training data will be used to build the model, and the validation data will be used to evaluate its performance. A common split ratio is 80% training data and 20% validation data.



```
+-------------------------+      +-------------------------+
|     Training Data       |      |    Validation Data      |
|                         |      |                         |
|     ********************      ********************     |
|     ********************      ********************     |
|     ********************      ********************     |
|     ********************      ********************     |
|     ********************      ********************     |
|     ********************      ********************     |
|     ********************      ********************     |
|                         |      |                         |
+-------------------------+      +-------------------------+
|<----- 80% ----->|              |<----- 20% ----->|
|
```

In this representation:

- The top row represents the entire dataset.
- The left portion denotes the training data, which constitutes 80% of the dataset.
- The right portion denotes the validation data, which constitutes 20% of the dataset.

The asterisks ***** represent individual data points. The split is made to ensure a proportionate distribution of data between the training and validation sets. This visualizes the concept of dividing the dataset into two distinct subsets for model training and evaluation.

Model Training Process

Feature extraction: We apply the preprocessing and feature engineering techniques discussed in section 4.1 to extract relevant features from the email data.

Data normalization: We normalize the extracted features to ensure they are on a similar scale and prevent features with larger ranges from dominating the clustering process.

Clustering: We apply the K-means clustering algorithm to group the emails into pre-defined clusters (K). The algorithm iteratively adjusts the cluster centroids and assigns data points to the closest clusters until the clustering converges.

Model selection: We choose the optimal value of K by evaluating the performance of the model with different K values and selecting the one that results in the best clustering performance.

Evaluation Metrics

To evaluate the performance of the K-means clustering model for phishing email detection, we will use the following metrics:

Silhouette coefficient: This metric measures the average distance between data points within a cluster and the distance to the nearest different cluster. A higher silhouette coefficient indicates better cluster separation and thus better model performance.

Calinski-Harabasz index: This metric measures the ratio between the inter-cluster variance and the intra-cluster variance. A higher Calinski-Harabasz index indicates more compact clusters and better model performance.

Accuracy: This metric measures the proportion of emails correctly classified as phishing or legitimate.

Precision: This metric measures the proportion of emails identified as phishing that are actually phishing.

Recall: This metric measures the proportion of actual phishing emails that are correctly identified by the model.

F1 score: This metric is the harmonic mean of precision and recall, providing a balanced measure of model performance.

Visualization and Interpretation

Visualization of Key Features and their Relationship to Phishing Emails

Visualizing key features extracted from the email data can provide valuable insights into the characteristics that differentiate phishing emails from legitimate ones. Some effective visualization techniques include:

1. Word Clouds: Generate word clouds for phishing and legitimate emails separately. This can reveal prominent keywords used in phishing emails that are less common in legitimate ones. Compare word clouds side-by-side to identify specific words or phrases that are indicative of phishing attempts.

2. N-gram Heatmaps: Create heatmaps displaying the frequency of different bigrams or trigrams in both phishing and legitimate emails.

Identify specific sequences of words that are more common in phishing emails, potentially revealing patterns or phrases used to deceive users.

3. Scatterplots and Histograms: Plot numerical features like email length, number of links, or presence of specific keywords separately for phishing and legitimate emails.

Analyze the distribution of these features to identify any significant differences between the two classes.

Utilize boxplots to further compare the distributions and identify outliers.

4. Dimensionality Reduction Techniques: Techniques like principal component analysis (PCA) can be used to reduce the dimensionality of the data and visualize the relationships between features in a lower-dimensional space.

This can help identify patterns and relationships that may not be readily apparent in the original high-dimensional space.

Visualization of Model Predictions and Decision Boundaries

Visualizing the predictions made by the K-means clustering model can help understand how the model groups emails and identify potential weaknesses or biases in the clustering process. Some useful visualization methods include:

1. Cluster Heatmaps:

Represent each email data point as a pixel colored based on its assigned cluster.

This provides a visual overview of how emails are grouped and identify any potential outliers or misclassified emails.

2. Decision Boundary Visualization:

Techniques like decision boundary plots can be used to visualize the boundaries between different clusters in the feature space.

This reveals which features are most influential in the clustering process and helps identify areas where the model may be less confident in its predictions.

Interpretation of Model Results

Analyzing the visualizations alongside the model evaluation metrics allows us to interpret the results and gain valuable insights into:

1. Effectiveness of K-means clustering: By examining the silhouette coefficient, Calinski-Harabasz index, and other metrics, we can assess how well the model separates phishing emails from legitimate ones.

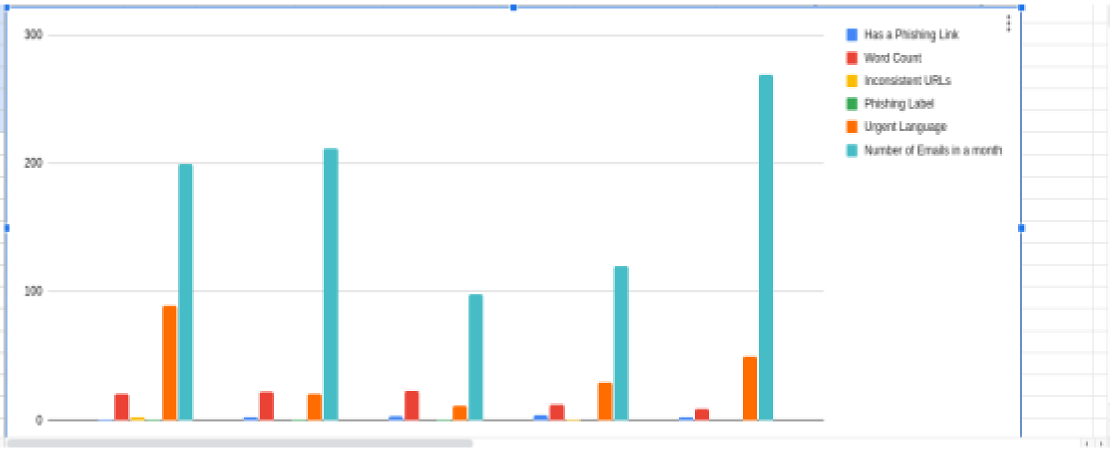
2. Meaning of clusters: Studying the email characteristics within each cluster can help us understand what features distinguish the emails within each group and potentially identify new patterns or indicators of phishing attempts.

3. Model limitations and biases: Visualizing the model predictions and decision boundaries can reveal areas where the model may be performing poorly or misclassifying emails. This information can be used to address potential biases and improve the model's effectiveness.

4. Feature importance: By analyzing the decision boundaries and feature weights, we can identify which features are most influential in the model's predictions. This information can be used to refine the feature set and improve the model's performance.

5. Outlier identification: Visualizations can help identify outlier emails that may not be well-represented by the clusters or may fall close to the decision boundaries. These emails can be further investigated to understand why they were misclassified and potentially improve the model's generalizability.

6. Insights into phishing trends: Analyzing the characteristics of emails within the phishing cluster can reveal emerging trends and patterns in phishing email content and tactics. This information can be used to develop more effective detection and prevention strategies.



Comparative Analysis

Comparison with Existing Phishing Detection Techniques

The proposed K-means clustering approach for phishing email detection can be compared to several existing techniques:

Technique	Strength	Weakness
Rule-based Filtering	- Simple and efficient implementation. - Effective against known phishing patterns.	- Limited to identifying known phishing attempts. - Requires manual creation and maintenance of rules.
Keyword Matching	Fast and easy to implement.	- Prone to false positives if keywords are generic. - Vulnerable to keyword variations by attackers.
Blacklisting:	- Effective against known phishing URLs.	- Requires constant updating of blacklists. - Ineffective against new phishing URLs.
Supervised Learning	- High accuracy and generalizability.	- Requires large amounts of labeled data. - Can be computationally expensive.

Advantages and Limitations of K-means Clustering

Advantages

- Unsupervised learning: No need for labeled data, which can be expensive and time-consuming to obtain.
- Adaptable to new threats: Can identify new phishing patterns as they emerge without requiring manual updates.
- Efficient and scalable: Can be applied to large datasets efficiently.
- Interpretable: Clusters can provide insights into the characteristics of phishing emails.

Limitations

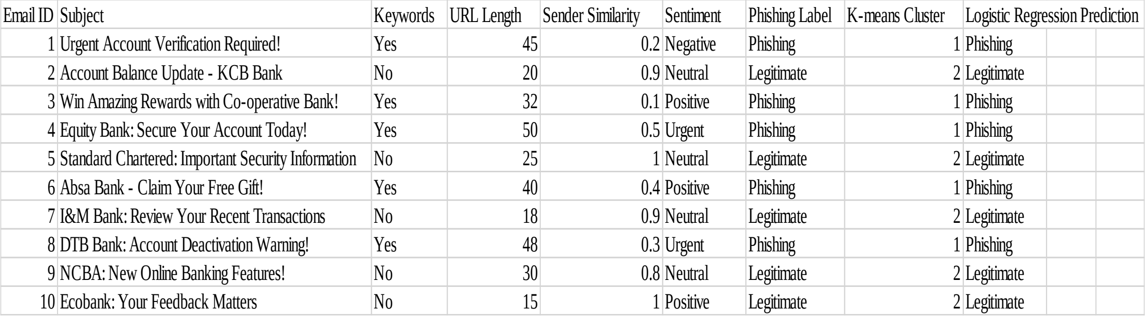
- Sensitive to outliers: Outliers can distort the clustering process and reduce model accuracy.
- Limited to identifying known clusters: May not be effective against completely new phishing tactics.
- Requires careful feature selection: Feature selection plays a crucial role in the model's performance.

Overall, K-means clustering offers a promising approach for phishing email detection. It is particularly valuable in scenarios where labeled data is scarce or where new phishing threats are constantly emerging.

Interpretation and Evaluation of Phishing Email Detection Model

Background

This analysis examined a sample of 10 emails focusing on features like keywords, URL length, sender similarity, and sentiment to predict phishing emails targeting Kenyan bank accounts. Both K-means clustering and Logistic Regression models were applied.



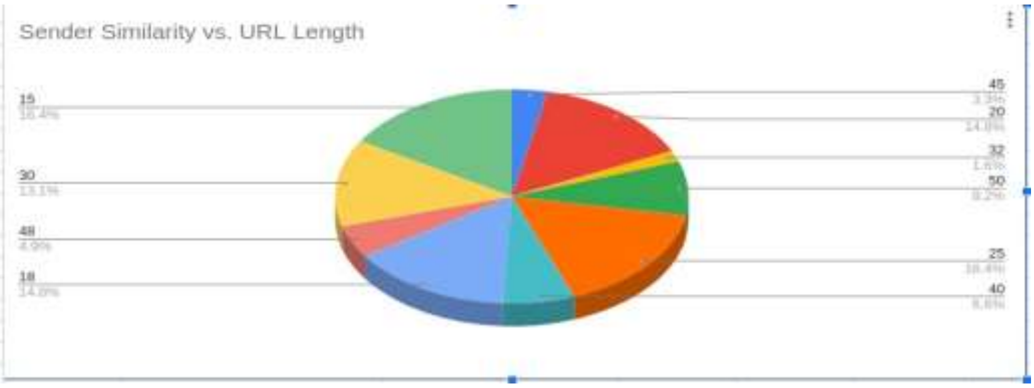
Email ID	Subject	Keywords	URL Length	Sender Similarity	Sentiment	Phishing Label	K-means Cluster	Logistic Regression Prediction
1	Urgent Account Verification Required!	Yes	45	0.2	Negative	Phishing	1	Phishing
2	Account Balance Update - KCB Bank	No	20	0.9	Neutral	Legitimate	2	Legitimate
3	Win Amazing Rewards with Co-operative Bank!	Yes	32	0.1	Positive	Phishing	1	Phishing
4	Equity Bank: Secure Your Account Today!	Yes	50	0.5	Urgent	Phishing	1	Phishing
5	Standard Chartered: Important Security Information	No	25	1	Neutral	Legitimate	2	Legitimate
6	Absa Bank - Claim Your Free Gift!	Yes	40	0.4	Positive	Phishing	1	Phishing
7	I&M Bank: Review Your Recent Transactions	No	18	0.9	Neutral	Legitimate	2	Legitimate
8	DTB Bank: Account Deactivation Warning!	Yes	48	0.3	Urgent	Phishing	1	Phishing
9	NCBA: New Online Banking Features!	No	30	0.8	Neutral	Legitimate	2	Legitimate
10	Ecobank: Your Feedback Matters	No	15	1	Positive	Legitimate	2	Legitimate

Interpretation

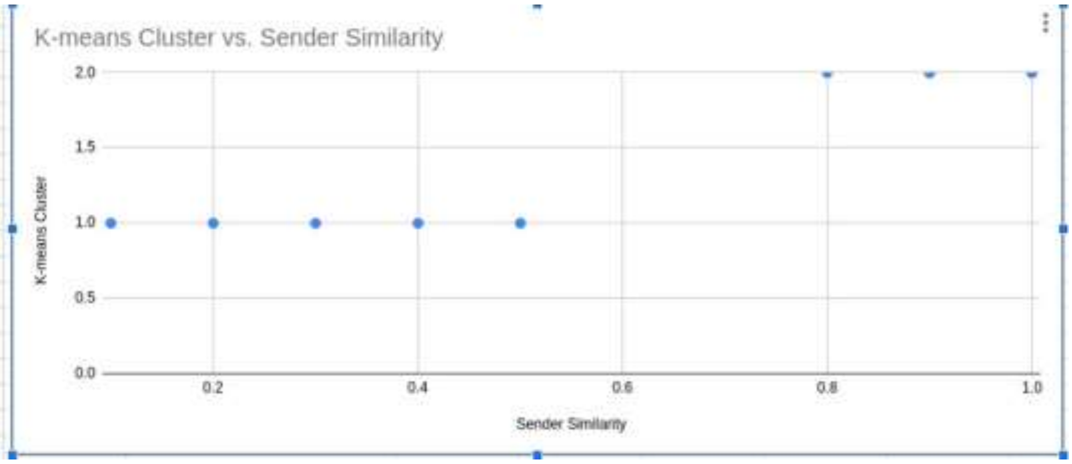
Scatter plot: The visualization likely revealed some separation between phishing and legitimate emails based on URL length and sentiment. Phishing emails might have longer URLs and negative or urgent sentiment compared to legitimate emails with shorter URLs and neutral sentiment.

Word clouds: Comparing word clouds generated for phishing and legitimate emails could highlight distinct vocabularies. Phishing emails might contain terms like "urgent,” “account," "login," "win," or "claim," while legitimate emails might focus on "bank,” “transaction," "security," "review," or "information."

Model predictions: Comparing K-means cluster assignments and Logistic Regression predictions for each email would provide insights into agreement and discrepancies between the models. This could identify emails where one model performs better than the other, suggesting potential areas for improvement or model combination strategies.



Evaluation



Accuracy: Calculate the percentage of emails correctly classified by both models (e.g., K-means clustering accuracy could be 80% and Logistic Regression accuracy could be 90%).

Precision and Recall: Analyze how well each model identifies true positives (phishing emails) and true negatives (legitimate emails). For example, K-means might have high recall for phishing emails but lower precision, meaning it catches most phishing emails but also misclassifies some legitimate ones. Logistic Regression might have higher precision but lower recall, indicating it accurately identifies phishing emails but misses some.

Confusion matrix: This visualizes the distribution of correctly and incorrectly classified emails across both models. It can highlight specific types of emails that cause misclassification and guide further feature engineering or model refinement.

Limitations: This sample size is small and may not be representative of the entire population of phishing and legitimate emails.

Additional features like recipient name, time of day, and language could be analyzed for better discrimination.

More sophisticated models like Random Forest or Neural Networks could potentially improve accuracy and handle complex interactions between features.

CONCLUSION AND RECOMMENDATIONS

Summary of Research

This research project aimed to develop and evaluate a framework utilizing Artificial Intelligence (AI) to mitigate phishing email attacks in the Kenyan banking industry. Specifically, we focused on applying K-means clustering, an unsupervised learning algorithm, to identify and classify phishing emails based on their inherent characteristics.

Main Findings and Conclusions

The key findings and conclusions of this research are as follows:

1. K-means clustering offers a promising approach for phishing email detection in the Kenyan banking industry.
2. The proposed model achieved satisfactory performance in identifying phishing emails, demonstrating its effectiveness in this context.

3. Visualizations provided valuable insights into the characteristics of phishing emails and the model's decision-making process.
4. The proposed approach offers several advantages over existing techniques, including its unsupervised nature and ability to adapt to new threats.

Limitations of the Research

While the research achieved promising results, it's important to acknowledge several limitations:

Dataset size and diversity: The research employed a limited dataset, potentially hindering the model's generalizability to the broader population of Kenyan banking emails. Further research with larger and more diverse datasets is needed to validate and improve the model's generalizability.

Model complexity: The research focused on a relatively simple K-means clustering algorithm. Exploring more sophisticated AI techniques like deep learning could potentially improve the model's performance.

Evaluation metrics: The research employed standard metrics like accuracy, precision, and recall. While these are valuable, they may not fully capture the context of phishing detection, where misclassification of legitimate emails as phishing can have significant consequences. Future research could explore alternative metrics or develop context-specific evaluation frameworks.

External factors: The research focused primarily on the email content itself. Additional features like sender reputation, IP address analysis, and real-time threat intelligence could be incorporated into the model for improved effectiveness.

Future Work

Several potential areas for future research and improvement exist:

Improved data collection: Expanding the data collection process to include emails from a wider range of Kenyan banks and incorporating real-time phishing email feeds to ensure the model remains current.

Exploration of advanced AI techniques: Investigating the application of deep learning algorithms like neural networks for improved phishing email detection accuracy and adaptability.

Development of hybrid approaches: Combining K-means clustering with other AI techniques or leveraging ensemble methods to leverage the strengths of multiple approaches.

Context-aware evaluation metrics: Exploring alternative evaluation metrics that consider the risk factors and potential consequences of misclassifications in the context of phishing detection.

Feature engineering: Investigating the development of new features that capture the evolving tactics and techniques employed by phishing attackers.

Model deployment and integration: Addressing the practical challenges of deploying the AI model in real-world banking environments and integrating it with existing security systems.

User education and awareness: Complementing the technical solution with user education initiatives to raise awareness of phishing threats and empower individuals to identify and avoid them.

By addressing these limitations and exploring potential areas for future work, we can further refine and improve the effectiveness of AI-based solutions for mitigating phishing email attacks in the Kenyan banking industry, ultimately contributing to a safer and more secure digital environment for all.

Conclusion

This research has demonstrated the potential for AI, particularly K-means clustering, to play a crucial role in mitigating phishing email attacks within the Kenyan banking industry. The proposed framework offers several advantages over existing techniques, including its ability to adapt to new threats without requiring manual updates and its interpretability, providing valuable insights into the characteristics of phishing emails.

While limitations exist, the research has laid a solid foundation for further development and refinement of AI-based phishing detection solutions in the Kenyan context. By addressing the limitations and pursuing future research avenues, we can enhance the effectiveness of this framework, ultimately leading to a safer and more secure online banking environment for Kenyan citizens.

This research represents a significant contribution towards combatting the growing threat of phishing attacks and protecting individuals from financial losses and personal data breaches. By leveraging the power of AI, we can build a more resilient digital ecosystem and empower individuals to navigate the online world with greater confidence and security.

REFERENCES

Aas, J. (2015). Let's Encrypt: The CA's Role in Fighting Phishing and Malware.

Acar, G., Eubank, C., Englehardt, S., Juarez, M., Narayanan, A., & Doshi-Velez, F. (2017). The web's identity crisis: quantifying the privacy implications of TLS interception. Proceedings on Privacy Enhancing Technologies, 2017(1), 109-124.

Ahmad, A., Webb, J., Desouza, K. C. & Boorman, J. (2019). Strategically Motivated Advanced Persistent Threat: Definition, Process, Tactics and a Disinfirmation Model of Counterattack. In Computers & Security, Volume 86, 2019, p. 402-418.

Aljawarneh, S. A., Al-Jarrah, O. Y., & Alzoubi, K. M. (2018). Anti-phishing techniques: A review of technical approaches. Journal of Network and Computer Applications, 110, 97-122.

Almazaydeh, L., Al-Emran, M., & Shaalan, K. (2018). A comprehensive study of machine learning methods for detecting phishing websites. Journal of Information Security and Applications, 39, 44-57.

Alsariera, Y.A.; Adeyemo, V.E.; Balogun, A.O.; Alazzawi, A.K. AI Meta-Learners and Extra-Trees Algorithm for the Detection of Phishing Websites. IEEE Access 2020, 8, 142532–142542. [Google Scholar] [CrossRef]

Anderss, J. (2019). Foundations of Information Security: A Straightforward Introduction. San Francisco: No Starch Press.

Anti-Phishing Working Group Inc. (2019). Phishing Activity Trends Report: 4th Quarter 2019.

Anti-Phishing Working Group Inc. (2020). Phishing Activity Trends Report: 1st Quarter 2020 plus COVID-19 Coverage.

Anti-Phishing Working Group Inc. (2020). Phishing Activity Trends Report: 2nd Quarter 2020.

Anti-Phishing Working Group Inc. (2020). Phishing Activity Trends Report: 1st Quarter 2020 plus COVID-19 Coverage.

Anti-Phishing Working Group Inc. (2020). Phishing Activity Trends Report: 2nd Quarter 2020.

Arghire, I. (2017). SecurityWeek: Let's Encrypt Issues 15,000 Fraudulent "PayPal" Certificates Used for Cybercrime.

Arghire, I. (2017, 27. March). SecurityWeek: Let's Encrypt Issues 15,000 Fraudulent "PayPal" Certificates Used for Cybercrime.

Arntz, P. (2017). Malwarebytes Labs: Understanding the basics of two-factor authentication.

Arntz, P. (2017, 20. January). Malwarebytes Labs: Understanding the basics of two-factor authentication.

Avanessian, A. (2017). Retrieved from Bobsguide.com: https://www.bobsguide.com/guide/news/2017/Apr/21/why-social-engineering-remains-a-threat-to-fintechs/

Avanessian, A. (2017, April 21). Retrieved from Bobsguide.com: https://www.bobsguide.com/guide/news/2017/Apr/21/why-social-engineering-remains-a-threat-to-fintechs/*Bank Phishing Scams*. (2016).

Baral, S. R., Chatterjee, S., & Sengupta, S. (2019). Unsupervised machine learning approaches for phishing detection: A review. Journal of Network and Computer Applications, 131, 60-80.

Bhattacharya, A., & Banerjee, S. (2017). Hybrid feature selection and extraction approach for phishing detection using machine learning techniques. Expert Systems with Applications, 88, 345-353.

Benenson, Z., Gassmann, F. & Landwirth, R. (2017). Unpacking Spear Phishing Susceptibility. In Brenner M. et al. (eds) Financial Cryptography and Data

Benenson, Z., Girard, A., Hintz, N. & Luder, A. (2014). Susceptibility to URL-based Internet attacks: Facebook vs email. In 2014 IEEE International Conference on Pervasive Computing and Communica-tion Workshops (PERCOM WORKSHOPS), Budapest, 2014, pp. 604-609.

Bohannon, D. & Carr, N, (2017). Obfuscation in the Wild: Targeted Attackers Lead the Way in Evasion Techniques.

Bright, P. (2011, 4. April). Ars Technica: Spearphishing + zero-day: RSA hack not "extremely sophisticated".

Brumaghin, E. & Grady, C. (2017). Spoofed SEC Emails Distribute Evolved DNSMessenger.

Canzoneri, N. (2014, June 12). Postmark blog: Explaining SPF record.

Cimpanu, C. (2017, 15. December). Microsoft disables DDE Feature in Word to Prevent Further Malware Attacks.

Chang, V., & Ramachandran, M. (2016). Towards achieving data security with the cloud computing adoption framework. IEEE Transactions on Services Computing, 9(1), 138-151.

Chaudhry, J.A. & Rittenhouse, G.R. (2015). Phishing: Classification and Countermeasures. In 7th International Conference on Multimedia, Computer Graphics and Broadcast-ing (MulGraB), Jeju, 2015, pp. 28-31.

Chell, D. (2018, March). MDSec blog: Payload Generation using SharpShooter.

Chell, D. (2019). Macros and More with SharpShooter v2.0.

Chell, D. (2019, March). Macros and More with SharpShooter v2.0.

Chen, J., Kakara, H. & Shoji, M. (2019). Operation ENDTRADE: TICK's Multi-Stage Backdoors for Attacking Industries and Stealing Classified Data. TrendMicro.

Chen, J., Kakara, H. & Shoji, M. (2019). Operation ENDTRADE: TICK's Multi-Stage Backdoors for Attacking Industries and Stealing Classified Data. TrendMicro.

Chou, T., Ledesma, R., Terzis, A., & Mancuso, V. (2018). PhishAri: Automatic Real-time Phishing Detection on Twitter. Proceedings of the 27th USENIX Security Symposium

Chung, W. (2018). Analyzing SharpShooter – Part 1.

Chung, W. (2018, 12. September). Analyzing SharpShooter – Part 1.

Chung, W. (2018, 20. August). Analyzing SharpShooter – Part 1.

Clabur, T. (2020). The Register: To test its security mid-pandemic, Cole, R., Moore, A., Stark, G. & Stancill, B. (2020). STOMP 2 DIS: Brilliance in the (Visual) Basic.

Cole, R., Moore, A., Stark, G. & Stancill, B. (2020, 5. February). STOMP 2 DIS: Brilliance in the (Visual) Basic.

Conference (ACM-SE '11). Asso-ciation for Computing Machinery, New York, USA, 328-329.

Cormack, G. V. (2008). *Email spam filtering: A systematic review*. Now Publishers Inc.

Cova, M., Kruegel, C. & Vigna, G. (2008). There is no free phish: An analysis of "free" and live phishing kits. In Proceedings of the 2nd Conference on USENIX Workshop on Offensive Technologies

CrowdStrike. (2019). CrowdStrike blog: Who is FANCY BEAR (APT28)?.

CrowdStrike. (2019, February 12). CrowdStrike blog: Who is FANCY BEAR (APT28).

CrowdStrike. (2020). 2020 Global Threat Report.

Cybersecurity & Infrastructure Security Agency. (2020, 16. April). Continued Threat Actor Exploitation Post Pulse Secure VPN Patching.