JCTA Journal of Computing Theories and Applications ISSN:3024-9104

Research Article

The Use of AI to Analyze Social Media Attacks for Predictive Analytics

Temitope Samson Adekunle[1], Oluwaseyi Omotayo Alabi[2,*], Morolake Oladayo Lawrence[3], Godwin Nse Ebong[4], Grace Oluwamayowa Ajiboye[5], and Temitope Abiodun Bamisaye[6]

[1] Computer Science Department, Colorado State University, United States;
e-mail : temitope.adekunle@colostate.edu
[2] Mechanical Engineering Department, Lead City University, Ibadan, Nigeria;
e-mail : alabi.oluwaseyi@lcu.edu.ng
[3] Computer Science Department, Baze University, Abuja, Nigeria;
e-mail : morolake.lawrence@bazeuniversity.edu.ng
[4] Data Science Department, University of Salford, United Kingdom;
e-mail : g.n.ebong@edu.salford.ac.uk
[5] Computer Science Department, Precious Cornerstone University, Ibadan, Nigeria;
e-mail : ajiboyegrace1910@gmail.com
[6] Computer Science Department, National Open University of Nigeria, Abuja, Nigeria;
e-mail : bamisaye9999@gmail.com
* Corresponding Author : Oluwaseyi Omotayo Alabi

Abstract: Social engineering (SE) presents weaknesses that are difficult to quantify in penetration testing directly. The majority of expert social engineers utilize phishing and adware tactics to convince victims to provide information voluntarily. SE in social media has a similar structural layout to regular postings but has a malevolent intrinsic purpose. Recurrent Neural Network-Long Short-Term Memory (RNN-LSTM) was used to train a novel SE model to recognize covert SE threats in communications on social networks. The dataset includes various posts, including text, images, and videos. It was compiled over a period of several months. Then carefully curated to ensure that it is representative of the types of content that are typically posted on social media. First, using domain heuristics, the social engineering assaults detection (SEAD) pipeline is intended to weed out social posts with malevolent intent. After tokenizing each social media post into sentences, each post is examined using a sentiment analyzer to determine whether it is a training data normal or an abnormality. Subsequently, an RNN-LSTM model is trained to detect five categories of social engineering assaults, some of which may involve information-gathering signals. Thus, the proposed SEA model yielded a classification precision of 0.82 and a recall of 0.79.

Keywords: Artificial Neural Network; Cybersecurity; Machine Learning; Random Forest Classifier; Social Engineering Attack.

Received: February, 7th 2024
Revised: March, 18th 2024
Accepted: March, 23th 2024
Published: March, 24th 2024

Copyright: © 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Social media platforms have become an important part of daily life for many people, providing a way to connect with others and share information. However, these platforms have also become a tool for spreading misinformation and attacking others. In recent years, there has been a growing interest in using AI to analyze social media data for predictive analytics. This can help identify patterns in online attacks and develop strategies for combating them. Facebook now has more users than any other social media platform, receiving billions of visitors daily. Additionally, during the pandemic, social media usage for online commerce and communication during physical restrictions increased dramatically. The increased social media usage encourages hackers to utilize security flaws to steal user information [1]. Social media is people-focused therefore "hacking" the system entails applying social engineering to take advantage of human aspects. One common technique is to pose as peers or bots in chat boxes to get private data [2]. Moreover, any hacker with

sufficient skills can communicate with anyone on the planet without the website administrator's consent. For instance, hackers may send spam communications to users while pretending to be banks in order to obtain their passwords or bank accounts. Furthermore, hackers can now simply track the activity of real users on social media networks by making straightforward Application Programming Interface API requests. Reconnaissance is a common first step in SE assaults [3]. Before initiating vicious attacks that sound plausible to the victims, the attacker spends much time researching user behaviors, such as their preferred products and routines [4], [5].

Social engineering attacks that target users' moral fallibility are unique because they exploit specific behavioral vulnerabilities in their targets. For example, users who are nervous about succeeding, afraid of taking control, or afraid of failure may be more likely to act rashly and fall victim to these attacks. Understanding these vulnerabilities makes it possible to develop better defenses against social engineering attacks. According to Symantec Security Response, just 4% of cyber-attacks are brought on by technical flaws and software exploitation methods [6]. Although our findings show that some measures are being taken to protect social media data, the vast majority of the analyzed posts indicated a need for further improvement in data security. This includes the security of both the social media platforms themselves as well as the applications and devices used to access them. Without these additional measures, users' data remains vulnerable to various threats, including unauthorized access, data breaches, and malware. Ultimately, the current state of social media security leaves much to be desired. No matter the machine type or operating system, network security only stops a small number of threats [7].

HBGary disregarded the Content Management System CMS flaw that causes unauthorized shell access because of incorrect Secure Shell SSH configurations, despite the security being otherwise quite conventional and simple. The issue is caused by carelessness, a common human error brought on by exhaustion or inexperience [8]. This integration has been used to develop features such as traffic filtering and intrusion detection, which can help improve social media data security. SDN allows for centralized network control, while Cisco DNA uses automation and analytics to optimize network performance. Together, these technologies can help to ensure that social media data is protected from unauthorized access and malicious attacks [9], [10]. To our knowledge, no study has yet been done that employs ML to categorize SE threats because of these entities' subjectivity. The data types needed for SE make acquiring datasets more challenging due to security concerns. Instead of a packet datagram unit, social media posts in our situation are primarily texts written in many languages [11], [12]. Because human characteristics and behavior are continually changing on social media, data collection must be ongoing (rather than done in stages) [13], [14]. Natural Language Processing (NLP) is also necessary for analyzing social media data, since it can process language related to human factors such as fear, anxiety, and other emotions. NLP can analyze and interpret unstructured text data to extract relevant information and predict user behavior. This is crucial for detecting and preventing social engineering attacks.

Artificial neural networks are frequently used in computer networking for threat detection [15]. Staudemeyer [16] suggests enhancing the classification accuracy of network threats by utilizing network traffic techniques and making the entire process of known harmful activity for detecting assaults. They build a (neural) network with two cells each in each of the four memory blocks. The authors' experimental results showed that the proposed Long Short-Term Memory (LSTM) model outperformed existing methods because it can track and correlate the continuous communication records over time. Similarly, in our study, we can train LSTM on the phrases in social media posts by treating the sequence of individual words as a time-step sequence. This allows LSTM to learn the underlying patterns in the data and make predictions about the sentiment of the posts.

Meanwhile, an RNN for intrusion detection was created by Krishnan and Raajan [17]. While using machine learning in routing technologies like SDN and Cisco DNA can provide significant benefits, simpler approaches can be used to gather threat intelligence. One such approach is to use the Simple Network Management Protocol (SNMP) to monitor network activity and identify potential threats. Another approach is using a random forest classifier, a machine learning algorithm that can identify patterns in large datasets. These simpler approaches may be less complex than integrating machine learning into routing technologies, but they can still provide valuable insights into potential threats. Despite working with big datasets, the suggested RNN classifies comparable threats more precisely and trains more

quickly. Similarly, Wu et al. [18] using the Network Security Laboratory Knowledge Discovery in Databases NSL-KDD dataset, another RNN-LSTM for Intrusion Detection System IDS was trained, and its accuracy was compared to that of Subversion SVN, Artificial Neural Network ANN, and Vinayakumar et al.[19]. This improvement gain was later shown to be feasible because LSTM overcomes the vanishing gradient drop and fixes the long-term dependency problem when training network data [20], [21]. The upgraded leNet-5 and LSTM neural network structures were directly merged to describe network threats' spatial and temporal cues. Deep learning for threat intelligence has long inspired cybersecurity researchers, but these models cannot identify social engineering assaults without network parameters. Instead, semantic sentences, a network and NLP domain hybrid, profiles Search Engine Advertising SEA disguising as social media posts.

In this study, we detect specific SE attack modifications in social media postings. We train an RNN-LSTM model. The datasets provided by the Social Computing Data Repository, SNAP, and Network Repository are all based on older services, and the speed of tweets makes it difficult to obtain enough data to provide meaningful context. Therefore, we decided to crawl Facebook for social media comments instead. Once we had collected sufficient data, we developed a pipeline for data preprocessing specifically designed to detect social engineering attacks (SEAD). This pipeline allowed us to clean and prepare the data for analysis, which was essential for developing the ML model [22], [23]. To identify posts that suggest a malicious intent to gather information, the SEADS model uses a variety of variables, including keyword matching, provenance filtering, and pattern recognition. These variables are used to model the language patterns of the posts and assign each one a sentiment score. The model then classifies the posts as SE attacks based on these scores. By analyzing the spatial-spectral language patterns of the posts, the model can detect and flag malicious content more accurately. Traditionally, social media analysis has focused on identifying and mitigating threats like cyberbullying, hate speech, or malicious content using ML techniques. However, the novel aspect lies in implementing deep learning methodologies for attack classifications on social media platforms. This innovative approach involves leveraging the multi-layered neural networks' capabilities to discern more intricate patterns within textual, visual, and contextual data, allowing for a more nuanced and accurate classification of various types of attacks. The main problem addressed by this research is the lack of effective methods for detecting and preventing social media attacks using AI. While traditional security measures, such as firewalls and antivirus software, are effective for detecting known attacks, they cannot detect novel or zero-day attacks. In addition, the sheer volume of data generated by social media platforms makes it difficult to manually analyze for patterns and trends that could be used to identify potential attacks. The gap in the literature is the lack of AI-based systems that can automatically analyze large volumes of social media data to identify potential attacks.

2. Method

In Figure 1, we visually represent how our SEAD tool detects potentially malicious social media posts. The process starts with crawling data from Facebook and collecting a large dataset of social media posts. Then, the data is preprocessed using natural language processing (NLP) and data cleaning techniques. Next, the data is labeled using machine learning algorithms, and a classification model is trained on the labeled data. Finally, the trained model detects malicious social media posts in real-time. As seen in the figure, the tool uses a combination of NLP and ML techniques to identify posts that may be intended to deceive or manipulate users. Our definition of malevolent includes pretexting, accusatory, and imperative behavior. First, Spyder is used to trawl demographic information from individual Facebook accounts and social media postings from the open posts of random individuals. Then, a recognizer for entities is developed to separate the perpetrator, target victim, and assault target the three primary entities from text-based posts. To categorize texts into predetermined categories such as people, places, organizations, everyday items (digital), device kinds, and actions, entity recognition uses the Natural Language Toolkit NLTK and SpaCy framework. For instance, "It was posted on Facebook that "Public Bank Customer Care has noticed a change in the password for your user account Seyi." The tuplet's three essential parts formed by the social media posts are "subject: customer service, victim: Seyi, and target: passwords."

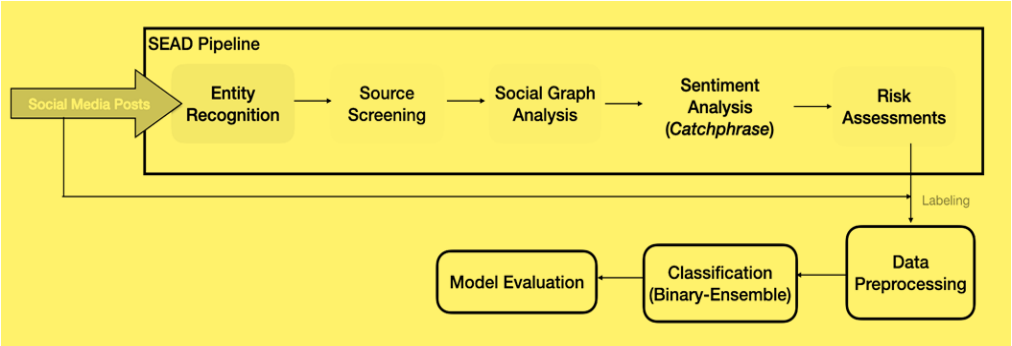

Figure 1. The Pipeline for Social Media Engineering Attack Classifications [23].

2.1. Data Input Analysis

SEAD functions on the presumption of guilt unless innocence is established. Using subject screening and filtering methods like Tesseract, SEAD is geared to ban SEA threats immediately. The input data analysis is comparable to stateful firewalls' blocking of signals. A building blacklist database's user accounts and known IP addresses of criminal people and botnets are compared to the previously recognized subjects and perpetrators. Input data analysis, unlike firewalls, searches for potential dangers in the application layer to spot hacked accounts with malicious intent. In addition to examining network headers like IP addresses, our model also considers the directionality of conversations, as illustrated in Figure 2. This means that a malicious source replying to a legitimate user who started the conversation is less likely to be flagged as suspicious. This is because the original message from the legitimate user is already part of the conversation and is, therefore, less likely to be malicious. This approach improves the model's accuracy and helps reduce false positives.

Figure 2. Calculate the maliciousness index of social media posts depending on the interaction state. A legitimate post in (a) must ask about previous encounters. If there haven't been any past interactions between the circles, a post with similar semantics in (b) gets red-flagged [23].

2.2 SE Detection Method based

Measurement methods frequently include sentiment analysis. To construct a sentiment analyzer that can identify SE assaults on social media posts, we use Google Auto ML. Because post models now express positive sentiment using positive adjectives and negative sentiment using negative adjectives, we need to develop a special sentiment analysis model [10], [24]. For instance, the present sentiment model will not detect the phrase "borrowing your account for emergencies" as a SE assault, even though it should. SEAD trains for keywords with malicious intentions using a bespoke Name-based entity recognition (NER). There are two steps at the core of NER. The NER initially looks for a word(s) that make up an entity. It's usual practice to tag entities with an inside-outside-beginning to denote their beginning and end [22], [25]. NER then classifies the identified entities into significant categories, such as person, organization, location, and in our instance, activities that suggest hostile intents. Two

experts assign a morphology-based label of "0" or "1" to dataset of instant each post, Label "0" denotes a benign social media message, such as "ideal conditions for hanging out," while label "1" denotes a potential SE assault, such as "steal your account."

2.3 Data Labeling and Risk Analysis

SEAD determines SEA's "integrity" across all social media post by averaging the threat factor across three detection components when doing risk analysis [3], [26]–[28]. Based on heuristics, each individual component is initially graded on a scale from 0 to 1. Each of the three factors is given equal weight. When a social media post's integrity has been compromised, such as when an account uploaded it or contained references to other accounts with verified identities, it is marked as true (1) during source screening a bad reputation, or vice versa. In the meantime, the post is marked as true (1) in the social graph to denote an indirect post, which includes postings that do not respond to prior interactions or mentions from unrelated personal and professional accounts. Last but not least, the sentiment analysis determines the sentiment score for each article based on professional keywords (One for positive and zero for negative). SEAD assesses whether a post has SEA elements based on these combined scores; a score of 0.5 is deemed safe (0) while a value of >0.5 is deemed hostile (1).

3. Results And Discussion

3.1 Datasets and Attack Classes

A Recurrent Neural Network-Long Short-Term Memory (RNN-LSTM) model is developed to detect the category of Social Engineering Attacks (SEAs) in social media posts, based on the linguistic characteristics of the posts. In addition, a machine learning model is used to classify the different types of SEA threats that are present in social media posts. This is because the language used in SEAs can be predictable and can be detected by analyzing the specific linguistic features of the posts. Table 1 provides examples of how the risk of SEAs in social media posts can be analyzed based on the components of the SEA Detection (SEAD) model. This model identifies the presence of SEAs based on four components: Target, Threat, Compromise, and Benefit. For example, if the Target component is identified as "account information" and the Threat component is identified as "phishing," then the SEA will likely be an account phishing attack. By analyzing the different components of SEAs, it is possible to understand better and predict the risk of such attacks on social media. For the model training 5,000 Facebook posts was choose with risk analysis scores greater than 0.5. Five categories pretexting, phishing, scareware, clickbaits, and quid pro quo are assigned to the dataset by two annotators. To avoid data imbalance, each assault class has an equal 1000 instances. The reliability of the data sets is examined, and any label inconsistencies are debated and resolved by the experts in accordance with their consensus. The classes are described as follows:

- *Pretexting* - Posts on social media in which the author adopts the personas of coworkers, law law enforcement, banking, and tax officials, or other anyone in a position to know. The pretexter asks ostensibly required questions to confirm the victim's identification. They develop wordlists for password guessing and cracking to get the essential personal information.
- *Phishing* - Email and SMS communications delivered by attackers pretending to be from a reliable and trusted source are known as phishing scams. These tactics take use of the victim's interest or terror to cause an illogical response to allegations of stolen credit cards, leaked images, and other sentimental material. The majority of the time, the victims are tricked into opening infected attachments or clicking on links to nefarious websites.
- *Scareware* - To trick people into believing that their system or user accounts have been compromised, scareware masquerades as pop-up notifications while browsing. Users are duped, and as a defense, they install suggested anti-threat tools that frequently risk themselves. As opposed to phishing, scareware is more relevant to actual user activities and contexts, which deceives people and lets down their guard.
- *Click baits* - The victim is baited into falling into the social engineering trap by being shown something enticing. For instance, skillfully worded email subject lines, free music downloads, or gifts with surveys rewards are worthwhile and deserve a few clicks. While

some social engineering attempts may be obvious, such as free mp3s that contain malware or free wallpapers that contain cryptocurrency mining software, the incentives for these attacks often go beyond what is immediately apparent. Attackers may seek to steal personal information, access sensitive systems, or even manipulate public opinion. It is important to be aware of the wide range of potential incentives for social engineering attacks, as this can help to identify suspicious activity and prevent harm. When individuals encounter deals that seem too good to be real, clickbait frequently succeeds against the weaker defense.

- *Quid Pro Quo* - a social engineering technique in which the attacker tries to exchange information for a service. These attacks prey on human weaknesses like curiosity and worry and are directed at less tech-savvy individuals. For instance, when faced with technical problems, end customers are more inclined to comply with IT assistance requests and freely divulge credentials for speedy solutions. Working from home has become more common recently, but few security landscapes have been thoroughly researched to identify possible vulnerabilities. An effort to use social engineering to trade services for information. These assaults take the use of feelings like curiosity and worry to prey on less tech-savvy individuals. Although remote login is more frequently used these days for working from home, few security landscapes have been thoroughly investigated to identify potential threats.

In data preprocessing, the training data are preprocessed to reduce input noise. Firstly, each instance is truncated to 250 characters or 30 words. Then, the text's stop words and arbitrary digits are removed (except common digits like date, month, and years to preserve temporal context). The text is further transcoded into a common Unicode format that supports emojis commonly found in iOS, Android, and Windows social posts. Lastly, some famous Internet Slang words are reconstructed to restore their linguistic meaning. The final dataset contains formatted 5,000 social posts; we split them at a 7:3 ratio for model training and testing.

Table 1. Risk analysis of social media posts (Test: Train).

SEA Types	Training		Testing	
	Instance Count	Word Count	Instance Count	Word Count
Pretexting	810	10102	205	2356
Phishing	810	11978	205	2056
Scareware	810	8013	205	1985
Clickbaits	810	10010	205	2435
Quid Pro Quo	810	9284	205	2006

3.2. RNN-LSTM turning parameter for Social Engineering Classification

Traditional non-neural network classification methods treat individual words as separate inputs, leading to a lack of contextual understanding within complete sentences. This limitation becomes pronounced in social media posts, which can be lengthy and rich in contextual nuances. To address this, we employed a Long Short-Term Memory (LSTM) model for classification, focusing on linguistic semantics rather than relying solely on statistical probabilities. Our LSTM model is designed to process multiple-word strings, enabling it to grasp the holistic meaning of sentences. It comprises five layers: an embedded layer for word representation, three hidden layers for decoding character, string, word, and phrase meanings, and an output layer for classification (refer to Figure 3). The embedded layer initializes with 100-length vectors representing each word. The positioning of words within these vectors is determined by their contextual relationships with preceding and succeeding words. For instance, "money" may be closely associated with "bank account," while "spam" may relate more to "email account." Following the embedded layer, a dense layer with 128 neurons processes the vectorized social media posts, which are first converted into sequences of integers and then one-hot-encoded to ensure uniform length during forward/backward propagation. The LSTM model employs softmax activation and the Adam optimizer, well-suited for multiclass classification tasks due to its efficacy in handling sparse gradients and noisy data. We chose sparse categorical cross entropy as the loss function, given the mutually exclusive nature of some SEA categories, where each sample corresponds to a single class. We integrated a

dropout layer after the initial hidden layer to prevent overfitting and maintain network balance. Subsequently, an additional LSTM layer with 128 cells followed by dense layers guides the network's feature extraction process. Finally, the network converges at the output layer, which outputs five values representing distinct SEA categories, thus effectively leveraging LSTM's contextual understanding to accurately predict and classify social engineering attacks based on linguistic nuances within social media posts. As with RNNs, the hidden layer size determines the complexity of the features the LSTM can learn. Social engineering attack detection should be adjusted based on the dataset and the specific attack types you're interested in.

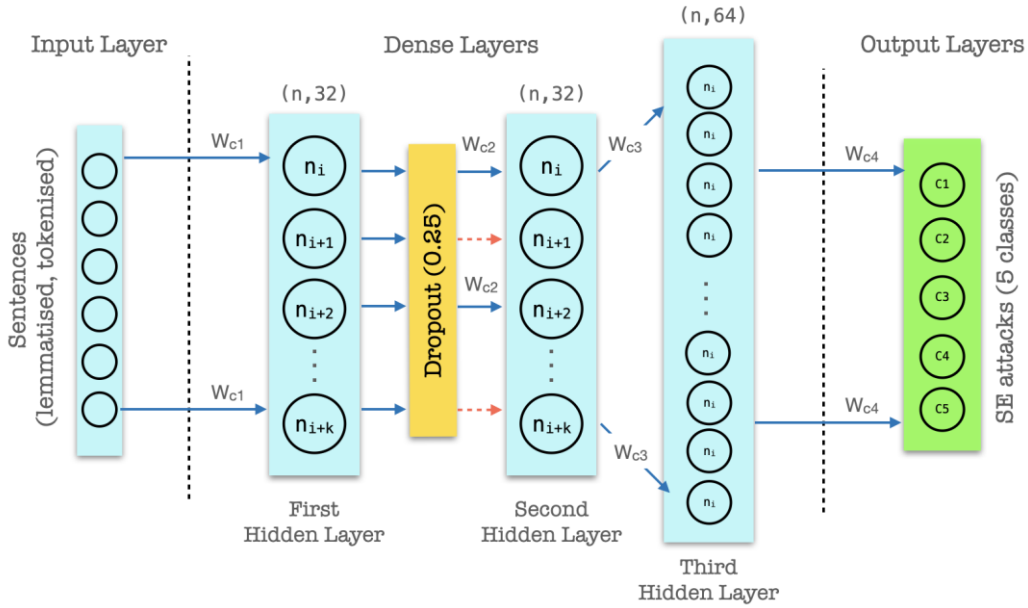

Figure 3. RNN-LSTM network architecture consisting of three hidden layers with dropout at 0.25

3.3 Performance Assessment

Since there is no standard benchmark to evaluate how well our model generalizes to new data, we tested its precision and recall against various well-known machine learning algorithms using a synthetic dataset. This dataset was designed to simulate real-world social media data and included features such as text, emojis, and other variables. The results showed that our LSTM model outperformed the other algorithms in terms of both precision and recall. This suggests that the model can generalize to new data and is robust to noise and variation in the data. Since common datasets like KDD Cup 99 and NSL-KDD don't have the necessary feature set, we employ 1,000 unseen samples that professionals have marked as the actual ground truth for model testing. Table 2 demonstrates that the proposed RNN-LSTM outperforms the other models on all measures, scoring 0.85 for precision and 0.80 for recall. The recall rate is generally slightly lower, which is typical for multiclass categorization of lengthy, unstructured texts. Longer sentences are difficult for traditional ML predictions because they are typically based on term frequency and a bag of terms. Surprisingly, despite being lighter and faster to train, typical ML-like KNN, DT, and RF hardly outperform neural networks in terms of performance. The decision tree (DT) algorithm is known to be effective at classifying text data when the data is simple and straightforward.

The k-nearest neighbor (KNN) algorithm is a clustering technique that can be used for classification without requiring large training datasets. Random forest (RF) is an ensemble technique that combines multiple decision trees to improve accuracy and generalization. In this case, we found that RF outperformed both DT and KNN, likely due to social media data's complex and varied nature. The disparity between PCA and DBN, on the other hand, is more severe since these algorithms classify words in a sentence as independent entities, losing certain spatial signals to the phrase's linguistic features. MLP, which is slightly less accurate than LSTM, also utilizes forward/backward propagation on a neural network to learn the meaning with the best NN settings and hyperparameters, we contrast an optimized LSTM

with an MLP. We ramify that sentence structure, including word choice and the relative order of occurrences, may include useful temporal information. Including the memory cell in the LSTM architecture propagates the error gradient at each learning level, promoting the desired behavior. While we cannot fully explain the inner workings of the neural network model, the ability to train the model on entire sentences rather than individual words gives it an advantage over traditional machine learning algorithms. However, it is important to note that neural networks are often considered "black box" models due to the difficulty of interpreting their inner workings When intentions are inferred from words, we conclude using LSTM that it is difficult to create a nearly perfect model. We must first consider the linguistic literacy gap when comparing intrinsic SEA intentions stated in words. Additionally, circumstances like timing, subjects, the criminal past of the author, the post's subject, and the political and cultural context of the participants are missing when SEA on social media is detected. In other words, certain social engineering attacks don't have verbal expressions and aren't ever represented by any linguistic semantics.

Table 2. A comparison of the SEA's Classification Precision and Recall for a number of well-known Classifiers.

Algorithm	Precision	Recall
DT(j47)	0.74	0.69
DBN	0.59	0.50
KNN	0.72	0.65
RF	0.80	0.74
PCA	0.53	0.44
RNN-LSTM	0.85	0.79

4. Conclusions

Social media posts are becoming targets for social engineering assaults (SEA). In order to trick victims into clicking on dangerous links and unwittingly disclosing critical information, they prey on their fears and insecurities. To lessen suspicion, attackers have recently become closer and more personal in their social media posts, making them sound like most other posts yet carrying an inherent motivation. Based on the SEAD pipeline is made to automatically categorize a social media post as harmful or legitimate based on source screening, social graph analysis, and sentiment analysis. We discover that the majority of SEA may be halted by closely examining postings made by dubious accounts at the source level. The LSTM model outperformed traditional machine learning algorithms, likely due to its ability to process whole sentences rather than individual words. This research is important for improving social media platform safety and helping users protect themselves from potential harm. Further research is needed to explore how these algorithms can be applied to real-world data and to understand the specific factors that lead to successful predictions.

Author Contributions Conceptualization, methodology: **Temitope Samson Adekunle**; software, validation, formal analysis, investigation: **Morolake Oladayo Lawrence and Godwin Nse Ebong**; resources, data curation: Grace Oluwamayowa Ajiboye; writing—original draft preparation, writing—review and editing, visualization, supervision, project administration **Oluwaseyi Omotayo Alabi**; funding acquisition: **Temitope Samson Adekunle, Morolake Oladayo Lawrence, Godwin Nse Ebong, Grace Oluwamayowa Ajiboye and Temitope Abiodun Bamisaye.**

Funding: This research received no external funding

Data Availability Statement: Data is unavailable due to privacy or ethical restrictions.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] H. Aldawood and G. Skinner, "Educating and Raising Awareness on Cyber Security Social Engineering: A Literature Review," no. December, pp. 62–68, 2019, doi: 10.1109/tale.2018.8615293.

[2] S. Tanwar, T. Paul, K. Singh, M. Joshi, and A. Rana, "Classification and Imapct of Cyber Threats in India: A review," in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Jun. 2020, pp. 129–135. doi: 10.1109/ICRITO48877.2020.9198024.

[3] T. O. Akande, O. O. Alabi, and J. B. Oyinloye, "A Review of Generative Models for 3D Vehicle Wheel Generation and Synthesis," *J. Comput. Theor. Appl.*, vol. 2, no. 2, pp. 148–168, Mar. 2024, doi: 10.62411/jcta.10125.

[4] S. A. Ajagbe and M. O. Adigun, "Deep learning techniques for detection and prediction of pandemic diseases: a systematic literature review," *Multimed. Tools Appl.*, vol. 83, no. 2, pp. 5893–5927, Jan. 2024, doi: 10.1007/s11042-023-15805-z.

[5] S. Dasgupta, A. Piplai, A. Kotal, and A. Joshi, "A Comparative Study of Deep Learning based Named Entity Recognition Algorithms for Cybersecurity," in *2020 IEEE International Conference on Big Data (Big Data)*, Dec. 2020, pp. 2596–2604. doi: 10.1109/BigData50022.2020.9378482.

[6] C. Lorenzen, R. Agrawal, and J. King, "Determining Viability of Deep Learning on Cybersecurity Log Analytics," in *2018 IEEE International Conference on Big Data (Big Data)*, Dec. 2018, no. April, pp. 4806–4811. doi: 10.1109/BigData.2018.8622165.

[7] D. Gumusbas, T. Yldrm, A. Genovese, and F. Scotti, "A Comprehensive Survey of Databases and Deep Learning Methods for Cybersecurity and Intrusion Detection Systems," *IEEE Syst. J.*, vol. 15, no. 2, pp. 1717–1731, Jun. 2021, doi: 10.1109/JSYST.2020.2992966.

[8] K. Shaukat, S. Luo, V. Varadharajan, I. A. Hameed, and M. Xu, "A Survey on Machine Learning Techniques for Cyber Security in the Last Decade," *IEEE Access*, vol. 8, pp. 222310–222354, 2020, doi: 10.1109/ACCESS.2020.3041951.

[9] Temitope S. Adekunle; Morolake O. Lawrence; Oluwaseyi O. Alabi; Adenrele A. Afolorunso; Godwin N. Ebong; Matthew A. Oladipupo, "Deep Learning for Plant Disease Detection," *Comput. Sci. Inf. Technol.*, vol. 5, no. 1, pp. 49–56, 2023, doi: 10.11591/csit.v5i1.pp49-56.

[10] A. Algarni, Y. Xu, and T. Chan, "An empirical study on the susceptibility to social engineering in social networking sites: the case of Facebook," *Eur. J. Inf. Syst.*, vol. 26, no. 6, pp. 661–687, Nov. 2017, doi: 10.1057/s41303-017-0057-y.

[11] T. Bakhshi, "Social engineering: Revisiting end-user awareness and susceptibility to classic attack vectors," in *2017 13th International Conference on Emerging Technologies (ICET)*, Dec. 2017, vol. 2018-Janua, pp. 1–6. doi: 10.1109/ICET.2017.8281653.

[12] R. Naidoo, "A multi-level influence model of COVID-19 themed cybercrime," *Eur. J. Inf. Syst.*, vol. 29, no. 3, pp. 306–321, May 2020, doi: 10.1080/0960085X.2020.1771222.

[13] A. Dan and S. Gupta, "Social Engineering Attack Detection and Data Protection Model (SEADDPM)," in *Advances in Intelligent Systems and Computing*, vol. 811, no. January, Springer Singapore, 2019, pp. 15–24. doi: 10.1007/978-981-13-1544-2_2.

[14] A. de Coning and F. Mouton, "Water Distribution Network Leak Detection Management," in *Proceedings of the 19th European Conference on Cyber Warfare*, Jun. 2020, vol. 2020-June, no. June, pp. 89–97. doi: 10.34190/EWS.20.088.

[15] D. Shafiei, S. A. Mostafavi, and S. J. Mehrabadi, "Geometrical optimization of city gate station's water bath indirect heater to minimization of fuel consumption," *J. Therm. Eng.*, vol. 9, no. 4, pp. 841–860, Aug. 2023, doi: 10.18186/thermal.1325287.

[16] J. S. Giboney, R. M. Schuetzler, and G. M. Grimes, "Developing a measure of adversarial thinking in social engineering scenarios," in *Proceedings of the 16th Pre-ICIS Workshop on Information Security and Privacy*, 2021, pp. 1–15.

[17] D. Berman, A. Buczak, J. Chavis, and C. Corbett, "A Survey of Deep Learning Methods for Cyber Security," *Information*, vol. 10, no. 4, p. 122, Apr. 2019, doi: 10.3390/info10040122.

[18] Y. Wu, D. Wei, and J. Feng, "Network Attacks Detection Methods Based on Deep Learning Techniques: A Survey," *Secur. Commun. Networks*, vol. 2020, pp. 1–17, Aug. 2020, doi: 10.1155/2020/8872923.

[19] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," *IEEE Access*, vol. 7, no. c, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.

[20] T.-T.-H. Le, J. Kim, and H. Kim, "An Effective Intrusion Detection Classifier Using Long Short-Term Memory with Gradient Descent Optimization," in *2017 International Conference on Platform Technology and Service (PlatCon)*, Feb. 2017, no. February, pp. 1–6. doi: 10.1109/PlatCon.2017.7883684.

[21] T. Zhang *et al.*, "Winglet design for vertical axis wind turbines based on a design of experiment and CFD approach," *Energy Convers. Manag.*, vol. 195, no. February, pp. 712–726, Sep. 2019, doi: 10.1016/j.enconman.2019.05.055.

[22] W. Alexan, E. Mamdouh, M. Elbeltagy, A. Ashraf, M. Moustafa, and H. Al-Qurashi, "Social Engineering and Technical Security Fusion," *Int. Telecommun. Conf. ITC-Egypt 2022 - Proc.*, no. August, 2022, doi: 10.1109/ITC-Egypt55520.2022.9855761.

[23] Y. Aun, M.-L. Gan, N. Haliza Binti Abdul Wahab, and G. Hock Guan, "Social Engineering Attack Classifications on Social Media Using Deep Learning," *Comput. Mater. Contin.*, vol. 74, no. 3, pp. 4917–4931, 2023, doi: 10.32604/cmc.2023.032373.

[24] A. Aljuhani and A. Alhubaishy, "Incorporating a Decision Support Approach within the Agile Mobile Application Development Process," in *2020 3rd International Conference on Computer Applications & Information Security (ICCAIS)*, Mar. 2020, pp. 1–6. doi: 10.1109/ICCAIS48893.2020.9096751.

[25] Z. Luo, W. Cai, Y. Li, and D. Peng, "The correlation between social tie and reciprocity in social media," in *Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology*, Aug. 2011, vol. 8, pp. 3909–3911. doi: 10.1109/EMEIT.2011.6023913.

[26] O. S. Ojo, M. O. Oyediran, B. J. Bamgbade, A. E. Adeniyi, G. N. Ebong, and S. A. Ajagbe, "Development of an Improved Convolutional Neural Network for an Automated Face Based University Attendance System," *ParadigmPlus*, vol. 4, no. 1, pp. 18–28, Apr. 2023, doi: 10.55969/paradigmplus.v4n1a2.

[27] S. A. Ajagbe, A. A. Adegun, A. B. Olanrewaju, J. B. Oladosu, and M. O. Adigun, "Performance investigation of two-stage detection techniques using traffic light detection dataset," *IAES Int. J. Artif. Intell.*, vol. 12, no. 4, p. 1909, Dec. 2023, doi: 10.11591/ijai.v12.i4.pp1909-1919.

[28] S. A. Ajagbe, O. A. Adeaga, O. O. Alabi, G. O. Ogunsiji, I. O. Oladejo, and M. O. Adigun, "An Alcohol Driver Detection System Examination Using Virtual Instruments," *J. Hunan Univ. Nat. Sci.*, vol. 50, no. 11, 2023, doi: 10.55463/issn.1674-2974.50.11.4.