SEBHA UNIVERSITY CONFERENCE PROCEEDINGS, VOL.03 NO. 3 2024
DOI: 10.51984/SUCP.V3I3.3748

المؤتمر السادس للعلوم الهندسية والتقنية
The Sixth Conference for Engineering Sciences and Technology (CEST-6)
Confrence Proceeeding homepage: https://cest.org.ly

Predicting Cyber Threats From Twitter Using Codeless LSTM Knime Model

*Inass A. Husien[a], Farij O. Ehtiba[a], Haitham Saleh Ben Abdelmula[b], Hend Abdelgader Eissa[c]

[a]Computer Science Department The Libyan Academy, Misurata,Misurata, Libya
[b]Computer Networks Department,College of Computer Technology, Zawia, Zawia, Libya
[c]Computer Technologies Department,College of Computer Technology Tripoli,Tripoli, Libya

Keywords:
Cyber Threats
CodelessApproach
Knime
LSTM
Twitter

ABSTRACT

Cybersecurity threats pose significant risks in the increasingly interconnected digital world. Traditional security measures struggle to keep pace with modern cyberattacks, necessitating innovative approaches for proactive threat detection. This paper explores a codeless approach using Long Short-Term Memory (LSTM) model within the Knime analytics platform to predict emerging cyber threats from Twitter data to avoid the complexity and hassle of writing and debugging code. Based on the implementation results, the proposed model achieved accuracy on the prediction around 74%.

التنبؤ بالتهديدات السيبرانية من تويتر بإستخدام نموذج LSTM Knime

*إيناس حسين[1] و فرج احطيبه[1] و هيثم بن عبدالمولى[2] و هند عيسى[3]

[1]قسم علوم الحاسب، أكاديمية الدراسات العليا مصراته، مصراته، ليبيا
[2]قسم شبكات الحاسوب، كلية تقنية الحاسوب الزاوية، الزاوية، ليبيا
[3]قسم تقنيات الحاسوب، كلية التقنية الالكترونية طرابلس، طرابلس، ليبيا

الكلمات المفتاحية:
التهديدات السبرانية
نهج بدون كود
نايم
LSTM
تويتر

الملخص

تشكل تهديدات الأمن السيبراني مخاطر كبيرة في العالم الرقمي بشكل متزايد. التدابير الأمنية التقليدية تكافح لمواكبة الهجمات الإلكترونية الحديثة، مما يستلزم أساليب مبتكرة للكشف الإستباقي عن التهديدات. تعمل هذه الورقة على إستخدام نموذج الذاكرة طويلة المدى (LSTM) ضمن منصة تحليلات Knime للتنبؤ بالتهديدات السيبرانية الناشئة من بيانات تويتر بأقل تعقيد وأفضل أداء. وبناء على نتائج التنفيذ فقد حقق النموذج المقترح دقة في التنبؤ بحوالي 74%.

1. Introduction

In the rapidly evolving digital landscape, the significance of cybersecurity cannot be overstated. As organizations increasingly rely on digital platforms for their operations and services, the risk of cyber threats has escalated exponentially [1]. This research paper delves into the critical domain of Cyber Security Intelligence (CTI), a proactive approach to managing and mitigating these threats.

Cyber Security Intelligence refers to the collection, analysis, and dissemination of information related to cyber threats, aimed at protecting an organization's digital assets [2]. It involves the use of advanced analytical tools and techniques to predict, identify, and respond to cyber threats, thereby ensuring the integrity, confidentiality, and availability of information systems.

The concept of CTI is rooted in the traditional intelligence cycle, but it is tailored to the unique challenges of the cyber realm [2]. It encompasses a wide range of activities, including threat intelligence, security analytics, incident response, and digital forensics. These activities are designed to provide organizations with a comprehensive understanding of their threat landscape, enabling them to make informed decisions about their cybersecurity strategies.

In the current era of data breaches, ransomware attacks, and cyber espionage, CTI has emerged as a crucial component of organizational resilience. By providing early warning of potential threats and enabling rapid response to incidents, CTI can help organizations to stay one step ahead of cybercriminals.

Cybersecurity threats have a profound impact on businesses, posing significant risks to their operations, finances, and reputation [1]. These threats encompass a wide range of malicious activities, including attempts to gain unauthorized access to sensitive information, manipulate data, extort money, or disrupt business operations. Cybercrime, which includes identity theft, malware attacks, online fraud, and more, can have devastating consequences for organizations

*Corresponding author:
E-mail addresses: Inass.husien@gmail.com, (F. O. Ehtiba)f.ehtiba@lam.edu.ly, (H. S. Ben Abdelmula)hsaa8383@gmail.com,
(H. A. Eiss) namarek2010@gmail.com

Article History : Received 02 March 2024 - Received in revised form 13 August 2024 - Accepted 21 October 2024

and individuals alike.

One key challenge faced by businesses is the inadequacy of traditional security approaches in addressing the evolving nature of cyber threats [1]. Conventional methods often lack scalability, exhibit slow response times, and struggle to detect advanced and insider threats effectively. This highlights the critical need for innovative research and technologies to develop more robust and comprehensive security measures that can defend against the growing diversity of network attacks.

Artificial intelligence (AI) has emerged as a powerful tool in combating cyber threats [1]. By leveraging AI-based cybersecurity techniques, businesses can enhance their defense mechanisms by automating threat detection, analyzing patterns in data to predict and prevent attacks, and improving incident response capabilities. AI technologies such as machine learning algorithms like the K-Nearest Neighbor (KNN) algorithm can be utilized to classify cyberattack patterns and enhance malware detection in sectors like finance.

Moreover, the integration of cybersecurity into business continuity planning is crucial for organizations to ensure operational resilience in the face of cyber incidents. Understanding common cyber threats, fostering a security culture within the organization, gaining boardroom support, and establishing a coordinated response are essential steps in bridging the gap between cybersecurity and business continuity. By investing in employee education and awareness programs, organizations can strengthen their cybersecurity posture and mitigate the human factor vulnerabilities that often lead to successful cyber-attacks.

Cyber Threat Intelligence (CTI) has emerged as a critical tool for organizations to stay ahead of these threats [2]. CTI refers to the process of collecting, analyzing, and disseminating information about potential or current cyber threats to enhance cybersecurity decision-making. It provides organizations with a comprehensive understanding of their threat landscape, enabling them to anticipate, detect, and respond to cyber threats more effectively.

Aiming to ease the cyber threat detection process and eventually make it faster too, this research paper proposes a codeless LSTM-based model for predicting emerging cyber threats from Twitter data using Knime. This could make the prediction process easier and simpler, and allow us to obtain accurate and reliable forecasts, and allow us to focus more on the data than the overhead of coding, contributing to the development of more proactive and effective cyber threat detection strategies.

2. Related Work

Several research studies have been conducted on the cyber threat prediction and proposed many approaches in order to achieve the security performance.

The authors in [3] presented a deep learning method for predicting DDoS attacks based on sentiment analysis of Twitter data. In this study, it uses a combination of a 13-layer CNN model and an improved LSTM model to extract features from Tweets, classify them into positive and negative classes, and then predict the occurrence of DDoS attacks based on the sentiment trends. Moreover, this research claims that most cyber-attacks are planned and discussed on social media platforms, and that negative sentiments can indicate a higher probability of an attack. It demonstrates the feasibility and effectiveness of using social media data as a source of information for cyber-security prediction and prevention.

In [4], the authors proposed a deep neural network architecture for processing cybersecurity information from Twitter. The CNN and BiLSTM models are also used to classify and extract information from tweets. Furthermore, the CNN classifier achieved an average true positive rate and true negative rate of 92%, while the BiLSTM NER model achieved an average F1-score of 92% in detecting specified labels. The paper also demonstrates the timeliness and relevance of the information extracted from Twitter, showing that some tweets mention vulnerabilities before they are disclosed in official databases and that some tweets reference zero-day exploits without identifiers.

While the research presented in [5] has proposed a deep learning approach for extracting cyber threat indicators from Twitter streams. The proposed framework is scalable, robust, and does not require any feature engineering. Moreover, it deploys various text representation methods, such as Keras embedding 1 and fastText 1. Keras embedding 1 performs better than the others for tweet data analysis. This study does not provide any justification or explanation for the choice of hyperparameters and network architectures for the deep learning models.

A novel Entity Recognition method for extracting cybersecurity-related entities such as attack techniques, vulnerabilities, and malware names, as proposed in [6], uses a Bidirectional Long Short-Term Memory (BiLSTM) network with a Conditional Random Field (CRF) layer to perform named entity recognition. The model, which they called XBiLSTM-CRF takes word embeddings and character embeddings as input, and outputs entity labels for each word. Furthermore, the proposed model achieves an F1-score of 88.7% on the test set. It shows that the model can recognize new and unseen entities, and handle complex and nested entities.

An implementation of NLP and CNN was applied to Tweets by Coyac-Torres et al in [7] for cyberattack detection. It aimed to both detect and classify four types of cyberattacks: phishing, spam, malware, and bot attacks. Moreover, the additional features extracted from the message structure and the URL structure have been proposed to enhance the performance of the CNN model. The proposed approach achieves an accuracy of 0.91 for cyberattack detection and 0.82 for cyberattack classification when deployed on real data. This model does not rely on specific characteristics of social networks, such as the number of followers, likes, or comments, to analyze the messages, making the proposed approach applicable to different sources of information, such as blogs, forums, and other social networks.

3. Cyber Threat Intelligence

Cyber Threat Intelligence (CTI) is crucial in the realm of cybersecurity, providing evidence-based knowledge about threat actors' motives, targets, and attack behaviors. This data is collected, processed, and analyzed to empower security teams to make informed decisions and shift from reactive to proactive security measures. Threat intelligence plays a vital role in shedding light on unknown threats, revealing adversarial motives and tactics, and enabling better understanding of threat actors' decision-making processes. It benefits organizations of all sizes by helping them process threat data effectively, respond faster to incidents, and proactively anticipate threat actors' next moves.

Cyber Threat Intelligence can be employed in predicting new threats and enhancing cybersecurity measures. By analyzing various properties such as threat actor skill, motivation, Tactics, Techniques, and Procedures (TT and P), and Indicators of Compromise (IoC), CTI provides valuable insights into potential cyber threats. Leveraging Machine Learning (ML) techniques with CTI allows for the analysis and prediction of threats based on these properties, enabling organizations to identify vulnerabilities and take appropriate control actions to improve overall cybersecurity [8].

Furthermore, the development of threat hunting models using ML algorithms has become essential in mitigating cyber-attacks. By implementing ML predictive analytics on datasets like OSTO-CID, organizations can utilize algorithms such as Decision Tree Classifier (DTC) to achieve high accuracy in identifying and responding to cyber threats effectively [9].

Artificial Intelligence (AI) also plays a significant role in enhancing information security in businesses. AI technologies can improve Machine Learning approaches to detect and prevent cyber-attacks by analyzing data from previous incidents, enabling behavior analysis, risk assessment, bot blocking, endpoint protection, and security task automation. While AI can assist cybersecurity experts in decision-making processes, it is essential to maintain a balance between risk and benefit as deploying AI may introduce new threats that need to be carefully managed [10].

In essence, cyber threat intelligence is not just about identifying threats but also about taking proactive measures to defend against them effectively [11]. By leveraging threat intelligence tools and platforms, organizations can enhance their security posture by staying ahead of evolving cyber threats and mitigating attacks efficiently [12].

A. Cyber Threat Intelligence Life Cycle

i. **Planning:** This is where goals have to be set for the intelligence program and make an understanding of what needs to be protected

and what could a cyber threat impact and generate the security requirements needed.

ii. **Data Collection**: This is where data that supports the objectives generated in the first step is collected. The threat data can be collected from different sources:
   - Threat intelligence feeds: these can be open source or commercial feeds that scrape social media, deep and dark webs for talks about new threats, aggregate cyberthreats news and keep track common attacks on Indicators of Compromise (IOCs).
   - Information-sharing communities: this includes forums, industry-specific sharing and analysis platforms and such communities where professionals and analysts worldwide exchange their knowledge, perspectives and their own threat intelligence.
   - Internal security logs: This provides cyberthreats and attacks records that organizations have been through.

iii. **Processing:** This is the process of filtering out the false positives, standardize and correlate the collected data and putting it in a usable format. Artificial intelligence (AI) and machine learning (ML) tools are used to automate this step that to spot trends and patterns in the data.

iv. **Analysis**: This is the process of generating intelligence information by the security analysts from the processed data to make recommendations that feed into the requirements that were put together in the first step.

v. **Dissemination:** This is where the recommendations and conclusions of the security analyst team are taken to the relevant stakeholders of the organization to discuss.

vi. **Feedback:** Finally, the stakeholders and the analysts ensure that the recent threat intelligence cycle reflects the requirements and if there were any gaps that should be covered in the next cycle.

B. Sources of Cyber Intelligence Data

One of the key challenges associated with CTI mining is the vast amount of data that needs to be collected and analyzed. This data can come from a variety of sources. The challenge is to collect and analyze this data in a way that provides actionable insights into the motives, targets, and attack behaviors of threat actors. Sharing threat intelligence is an essence of protecting organizations' security against fast emerging cyber threats. There are two types of cybersecurity information sources, formal sources such as the National Institute of Standards and Technology's (NIST) National Vulnerability Database (NVD) [13], Common Vulnerabilities and Exposures records (CVEs) [14], Mitre ATT&CK knowledge base [15], etc. and information security communities such as Facebook Threat Exchange and IOC Bucket, where professionals share threat intelligence in real time to help in protecting and preventing possible attacks [16-17].

In addition, there are various informal sources such as blogs, developer forums, chat rooms and social media platforms like X (formally known as Twitter) [18], Reddit [19] and Stack Overflow [20], including newspapers, magazines, video sharing sites, wikis and blogs. Such unstructured and publicly available sources are called Open-Source intelligence (OSINT) [21]. Web crawlers can be used to crawl the CTI data that's publicly available. Application Programming Interfaces (APIs) provided by a lot of social media platforms can additionally be used to collect and analyze threat information shared by professionals and organizations.

In a recent survey, 60% of respondents who have been victimized by cyber-attacks said the breach happened from not applying the patch to a known vulnerability, and 52% said not automating patch management process have put their organizations at a disadvantage [22]. Another report found that 75% of threats are first disclosed online giving a 7 days median delay between the time it is first reported online and the time a patch published [23]. This shows how vital patch and vulnerability management process can be and how important making the most of cyber threat data valuable knowledge.

4. Methodology

A. Data Preprocessing

The dataset used had more than 20,000 records of tweets that were collected throughout the year of 2018 related to cyber security events and threats. The owners [24] used a stream listener to filter tweets based on a set of keywords relevant to cyber security obtained the data set from. The data set contains both relevant and irrelevant tweets, as well as tweets belonging to different categories of cyber security events, such as vulnerability, ransomware, DDoS, data leak, etc.

The dataset was transformed to consist of only 3 features/attributes: Tweet ID, Text (of the tweet), and Type (threat detected or not). It was also cleaned off of duplicate and empty Tweets for it to finally be of 4,094 Tweets. The data set is split into 70% training set and the remaining 30% as test set. See Table 1 for an example of dataset entries.

Table 1: Dataset Records

ID	Text	Type
1. 5b8876f9b b325e65fa 2. 7e78e4	3. You can't get to courage without walking through vulnerability	4. 0
5. 5b8876f9 bb325e65 fa7e78e5	6. Data leak from Huazhu Hotels may affect 130 million customers	7. 1

B. Utilized Technologies

The algorithm that's going to be implemented, as mentioned above, is LSTM networks. For an easy quick implementation Knime software is used. Knime provides easy setup of a lot of deep learning algorithms with the creation of data workflows without the need for coding.

You will see in the next section how the workflow for this paper's methodology has been setup and implemented.

5. Implementation

In Knime, individual tasks within the data analysis workflow are represented by nodes. These nodes are like building blocks that connect to create a complete analysis pipeline. Our model's work flow components consist of the following node groups:

A. Reading and Preprocessing

a) **CSV Reader**: This node reads a CSV file from a given location, where the dataset file is kept, and outputs a table with the data.

b) **Partitioning**: This node splits the input table into two parts: the training set and the test set.

c) **Preprocess Training Set and Preprocess Test Set:** these are meta-nodes that contain sub-nodes for preprocessing the data, which demonstrated in Figures 1 and 2.
   1. **String to Document**: this step converts raw text strings into documents that can be processed by other nodes.
   2. **Text Preprocessing**: This step performs various operations on the documents, such as tokenization, normalization, filtering, and stemming.
   3. **Truncate:** This step truncates the documents to a fixed length of 32, discarding any tokens beyond that limit.
   4. **Dictionary Replacer**: This step replaces tokens in the documents with corresponding numerical codes from the created dictionary.
   5. **Zero Pad**: This step adds zeros to the end of the documents to make them all have the same length.
   6. **Create Collection Column**: This step creates a collection column from the documents, which can be used as input for other nodes, such as machine learning models.
   7. **Table Writer**: This step saves the preprocessed data and the dictionary to a file.
   8. **Unique Term Extractor**: Extracts unique terms from the training dictionary.
   9. **Joiner:** Joins the extracted terms with an index from the training dictionary.
   10. **Missing Value:** Handles missing values, defining default values for new terms not present in the training dictionary.
   11. **Number to String**: This node converts numeric values to string values, in this case it coverts the type column from integer to string, which is required for Keras Network Executer node.

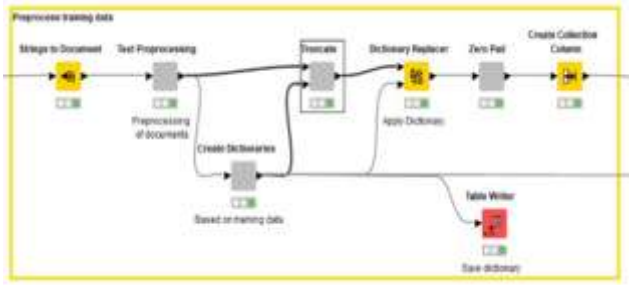

Fig. 1. Training set processing meta-node

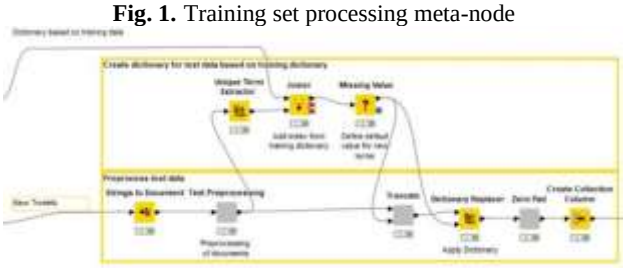

Fig. 2. Testing set processing

B. Define Network

a) **Keras Input Layer:** This is where the model receives its input data. This layer accepts sequences of varying lengths, up to a maximum length, which is 32 here. In our case, if the input data is a collection of tweets, each tweet may have a different number of words, but the input layer can handle them as long as they are not longer than the max sequence length.

b) **Keras Embedding Layer:** This layer transforms the input sequences into dense vectors of fixed size, which are trainable. It uses a word index lookup structure to understand the meaning of each token in its input. If the input data is a collection of tweets, each word in a tweet may be assigned a unique integer index, and the embedding layer will map each index to a vector representation that captures its semantic and syntactic features. The "+ 1" is for the special index 0, which is reserved for padding or unknown tokens.

c) **Keras LSTM Layer:** This is the core, the LSTM layer. "128 units" means there are 128 neurons or nodes in this layer. Each unit has an internal state that can store information over time, and can learn when to forget, update, or output its state based on the input and output sequences. In an input of a collection of tweets, the LSTM layer can learn the temporal dependencies and patterns among the words in each tweet, and generate a hidden state vector that summarizes the tweet's content and sentiment.

d) **Keras Dense Layer:** A dense layer is a kind of hidden layer where every node or neuron is connected to each node/neuron present in the previous and next layers, in this case, using binary classification. For example, here where the task is to classify Tweets as relevant or irrelevant, the dense layer can take the hidden state vector from the LSTM layer and apply a sigmoid function to produce a probability score between 0 and 1, where 0 means irrelevant and 1 means relevant.

C. Training and Prediction

a) **Keras Network Learner:** This node trains the neural network model on the training data, using a given optimizer, loss function, metrics, epochs, and batch size.

b) **Keras Network Executor:** This node uses the trained network model to make predictions on the test data, and outputs a table with the predicted values and the original values.

D. Evaluation

a) **Rule Engine:** This node applies rules to assign classes or manipulate data based on conditions, such as if then-else statements.

b) **Scorer:** This node evaluates the performance of the prediction model, by comparing the predicted values and the original values, and calculating various metrics, such as accuracy, precision, recall, F1-score, etc.

6. Results and Observations

This section details the performance metrics and observations from the model evaluation. Our evaluation focused on measuring the effectiveness of a Long Short-Term Memory (LSTM) model for classifying tweets as relevant or irrelevant to evolving cyber threats. The performance of our LSTM model was evaluated using a variety of metrics:

A. Confusion Matrix

Analysis of the confusion matrix revealed a higher number of false positives (133 irrelevant tweets classified as relevant) compared to false negatives (183 relevant tweets classified as irrelevant). This suggests the model might be overly cautious, classifying some ambiguous or noisy irrelevant tweets as relevant. The confusion matrix is shown in table 2:

Table 2: Confusion Matrix

Actual Classification	Predicted Classification: 0	Predicted Classification: 1
0	307	133
1	183	606

B. Precision and Recall

To further understand the model's performance, we computed precision and recall for both relevant and irrelevant tweet classifications. For the relevant class, precision was 0.763, indicating that 76.3% of the tweets classified as relevant were indeed relevant. Recall was 0.825, suggesting that 82.5% of all relevant tweets were correctly identified. For the irrelevant class, precision was 0.633, and recall was 0.541. These results demonstrate that while the model was moderately effective at identifying irrelevant tweets, there is still room for improvement in terms of both precision and recall.

C. F1-Score

To provide a more detailed analysis of the model's effectiveness, we computed the F1-scores for both relevant and irrelevant tweet classifications. The F1-score for the irrelevant class was 0.698, indicating a relatively high precision and recall in identifying tweets that are not related to cyber threats. For the relevant class, the F1-score was 0.768, reflecting the model's robust ability to correctly classify tweets pertinent to cyber threats.

D. Coehn's Kappa

To assess the model's inter-rater reliability, we calculated Cohen's Kappa. Cohen's Kappa is a statistical measure that quantifies the agreement between two raters (or, in this case, a human annotator and the model) on a categorical scale. It corrects for chance agreement, providing a more accurate assessment of the extent to which the two raters agree beyond what would be expected by chance.

Cohen's Kappa for our model was 0.378, which indicates a fair agreement between the model's predictions and human annotations. While this suggests some level of consistency, it also highlights the need for further improvements to enhance the model's reliability.

E. Receiver Operating Characteristic (ROC)

In addition to the previous metrics, we generated the ROC curve to further assess the model's classification ability. The ROC curve, depicted in Figure 3, illustrates the true positive rate (sensitivity) against the false positive rate (1specificity) at various threshold settings. The area under the curve (AUC) is 0.8134, which signifies a strong ability of the model to distinguish between relevant and irrelevant tweets. An AUC of 0.8134 indicates that the model has an 81.34% chance of correctly differentiating between the two classes.

F. Comparison

This section compares the performance of our LSTM model, implemented in KNIME with the LSTM model used by [3]. While both models employed the same algorithm, they differed in their datasets and implementation approaches. The results of our experiments are summarized in the following table:

Table 3: Results Comparison

Metric	Our Study (KNIME)	Aliguliyev and Abdullayeva [3]
Dataset	Collection of Cyber Threat Indicators in Twitter Stream	US Tweet Data
Precision	0.698	0.8865
Recall	0.683	0.7522
F1-Score	0.733	0.8138

As shown in the table 3, our model achieved 0.698 for precision, 0.683 for recall, and 0.733 for F1-score. While these results are lower than those reported by Aliguliyev and Abdullayeva [3], it's important to note that the datasets used in the two studies were distinct. The differences in data distribution, language patterns, and cyber threat characteristics could have influenced the model's performance.

Furthermore, the implementation details and hyperparameter settings employed in the two studies might have contributed to the observed differences. In [3] the authors used a code-based implementation, while our model was built using KNIME.

While code-based implementations offer greater flexibility and control, they can be time-consuming to develop and require specialized programming skills. KNIME, on the other hand, is a low-code/no-code platform that provides a visual interface for building and deploying machine learning models. This can reduce development time and make it easier for non-programmers to create and experiment with models.

Further investigation into these factors would be necessary to gain a deeper understanding of the reasons behind the observed performance differences.

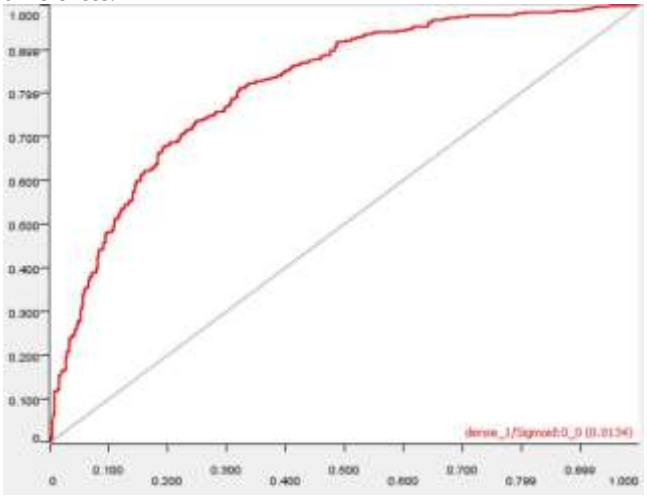

Fig. 3. ROC Curve

7. Conclusion and Recommendations

In this study, a KNIME workflow was used to build and evaluate a LSTM model for classifying tweets as relevant or irrelevant to cyber threats. An AUC of 81.34% was obtained through the model. The confusion matrix in Table 2 and the F1-scores indicate that the model has some difficulty in distinguishing between the two classes, and may be confused by some ambiguous or noisy tweets.

The model's performance is satisfactory, but not optimal. Further experimentation and tuning to achieve better results are suggested as follows:

1. Incorporating external knowledge and resources.
2. Extending the number and types of entities to be extracted from tweets, such as threat actors, attack vectors, or mitigation strategies, to provide more comprehensive and actionable threat intelligence.
3. Developing methods to measure and enhance the trustworthiness and credibility of the information extracted.
4. To improve performance, we could explore data augmentation techniques to enrich the training set with more relevant and irrelevant tweet examples. Additionally, tuning hyperparameters like epochs and batch size, or investigating more complex model architectures, could potentially lead to better accuracy.

8. References

[1]- Ahmad, A., Maulana, R., & Yassir, M. (2024). Cybersecurity Challenges In The Era Of Digital Transformation A Comprehensive Analysis Of Information Systems. Journal Informatic, Education and Management (JIEM), 6(1), 7-11.

[2]- Saeed, S., S. A. Suayyid, et al. (2023). "A Systematic Literature Review on Cyber Threat Intelligence for Organizational Cybersecurity Resilience." Sensors 23(16): 7273.

[3]- R. M. Aliguliyev, and F. J. Abdullayeva, “Deep Learning Method for Prediction of DDoS Attacks on Social Media,” Advances in Data Science and Adaptive Analysis, vol. 11, no. 01, Apr. 2019, doi: 10.1142/S2424922X19500025.

[4]- N. Dionísio, F. Alves, P. M. Ferreira, and A. Bessani, “Cyberthreat Detection from Twitter using Deep Neural Networks.” 2019.

[5]- S. K, P. Balakrishna, V. R, and S. KP, “Deep Learning Approach for Enhanced Cyber Threat Indicators in Twitter Stream.” 2020.

[6]- P. Ma, B. Jiang, Z. Lu, N. Li, and Z. Jiang, “Cybersecurity named entity recognition using bidirectional long short-term memory with conditional random fields,” Tsinghua Sci Technol, vol. 26, no. 3, pp. 259–265, Jun. 2021, doi: 10.26599/TST.2019.9010033.

[7]- J. E. Coyac-Torres, G. Sidorov, E. Aguirre-Anaya, and G. Hernández-Oregón, “Cyberattack Detection in Social Network Messages Based on Convolutional Neural Networks and NLP Techniques,” Mach Learn Knowl Extr, vol. 5, no. 3, pp. 1132–1148, Sep. 2023, doi: 10.3390/make5030058.

[8]- A. Yeboah-Ofori et al., “Cyber Threat Predictive Analytics for Improving Cyber Supply Chain Security,” IEEE Access, vol. 9, pp. 94318–94337,2021,[Online].Available: https://api.semanticscholar.org/ CorpusID :235826142

[9]- A. J. E. T., O. E. A., and A. G. A., “Development of Threat Hunting Model Using Machine Learning Algorithms for Cyber Attacks Mitigation,” 2022 International Conference on Computational Science and Computational Intelligence (CSCI), pp. 1010–1015, 2022.

[10]-S. A. Alawadhi, A. Zowayed, H. Abdulla, M. A. Khder, and B. J. A. Ali, “Impact of Artificial Intelligence on Information Security in Business,” in 2022 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS), IEEE,Jun.2022,pp.437–442.doi: 10.1109/ICETSIS55481.2022.9888871.

[11]-Cisco, “What Is Cyber Threat Intelligence?” [Online]. Available: https://www.cisco.com/c/en/us/products/security/ what-is-cyber-threat-intelligence.html

[12]-VMWare, “What is Threat Intelligence?” [Online]. Available: https://www.vmware.com/topics/glossary/content/threat-intelligence. html

[13]- “National vulnerability database.” [Online]. Available: https://nvd.nist.gov/vuln

[14]- “Common Vulnerabilities and Exposures records (CVEs).” [Online]. Available: https://www.cvedetails.com/

[15]- “Mitre ATT&CK knowledge base.” [Online]. Available: https://attack.mitre.org/

[16]- “Facebook ThreatExchange.” [Online]. Available: https://developers.facebook.com/products/threatexchange

[17]- “IOC bucket”, [Online]. Available: www.iocbucket.com/

[18]- “X (formally Twitter).” [Online]. Available: https://twitter.com/?lang=en

[19]- “Reddit.” [Online]. Available: https://reddit.com

[20]- “StackOverflow.” [Online]. Available: https://stackoverflow.com

[21]- “OSINT Framework.” [Online]. Available: https://osintframework.com/

[22]- “Costs and Consequences of Gaps in Vulnerability Response,” Ponemon Institute. [Online]. Available: https://www.servicenow.com/lpayr/ ponemonvulnerability-survey.html

[23]-Bill Ladd, “The Race Between Security Professionals and Adversaries.” [Online]. Available: https://www.recordedfuture.com/blog/vulnerabilit y-disclosure-delay

[24]-V. Behzadan, C. Aguirre, A. Bose, and W. Hsu, “Corpus and Deep Learning Classifier for Collection of Cyber Threat Indicators in Twitter Stream,” in 2018 IEEE International Conference on Big Data (Big Data), IEEE, Dec. 2018, pp. 5002–5007. doi: 10.1109/BigData.2018.8622506.