International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group| Volume 1, Issue 2, 37-45, 2020
ISSN: 3050-922X | https://doi.org/10.63282/3050-922X.IJERET-V1I2P105

Original Article

Deep Learning for Cyber Risk Management in Financial Services: A Case Study of Data Breach Prediction

Archana Pattabhi
Executive Leader in AI, Cybersecurity & Risk, SVP Citi; Member, Forbes Technology Council; CIO/CISO AdvisoryBoard, The Executive Initiative.

Abstract *- Amidst a rapidly growing and highly computerized environment, the financial services sector remains one of the most vulnerable to cyber threats and data breaches that have high operational and image risk implications. This paper aims to examine how deep learning can be used to identify factors that can be used in predicting data breaches, especially within the financial sector. With data from past security breaches and their metadata, the suggested deep learning model shall use LSTM and CNN networks in a multidimensional data set containing IT infrastructure data and records of user behaviour and threat intelligence feeds from third parties. We then determine the current cyber risk environment and explain why conventional risk management approaches are insufficient. We also explain how we managed the data and feature collection, derived the features from raw data, designed the architectures of the models, and evaluated the results. The performance of the proposed hybrid deep learning model has been tested on a benchmark dataset Contemplating real-world cyber incidents and compared to the results of classic machine learning methods like the Random Forests and Support Vector Machines (SVMs) where we have the numbers of 12% average of F1-score. This research is valuable because it proposes a model for measuring future cyber threats for SOC in the financial industry. Therefore, a proactive, artificial intelligence cybersecurity solution can have a prospective impact in lessening the blur incidence rate and enhancing compliance with the legislation.*

Keywords *- Cyber Risk Management, Data Breach, Deep Learning, Financial Services, LSTM, CNN, Threat Intelligence, Information Security, Risk Assessment*

1. Introduction

1.1 Importance of Deep Learning for Cyber Risk Management in Financial Services

In the context of the existing and constantly growing number of threats, conventional risk management methodologies prove irrelevant and insufficient for financial services. Such threats as APT, ransomware, and phishing require intelligence and dynamism to counter the attacks in today's world. [1-3] Deep learning can be a powerful solution to this problem because it presents superior instruments for cyber risk detection, prediction, and management compared to traditional approaches. As a result, in this section, we outline how deep learning applies to cyber risk management in the financial services industry under specific headings.

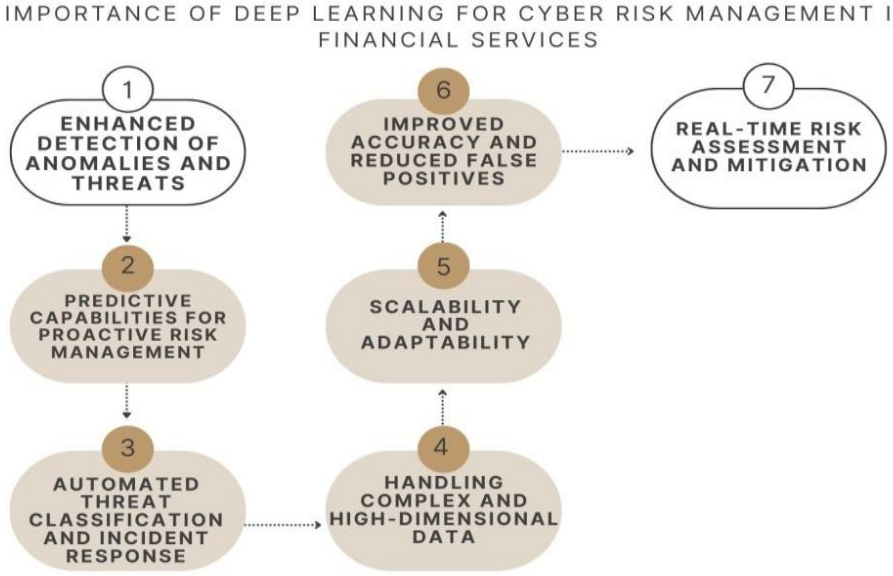

Fig 1: Importance of Deep Learning for Cyber Risk Management in Financial Services

- **Enhanced Detection of Anomalies and Threats:** Conventional or conventional cybersecurity techniques consist of rules or a dynamic signature-based approach. However, the authors pointed out that cybercriminals are growing more sophisticated, making it difficult to deal with them with simple rules. Some deep learning models like CNN and RNNs can learn from data and are able to detect patterns that conventional systems might not easily notice. These models can take a lot of transactional data, usage logs, and network traffic and make out for shifts from the normal traffic patterns, such as time anomalies, data flow anomalies, and login anomalies in the network.
- **Predictive Capabilities for Proactive Risk Management:** Thus, deep learning is used to manage cybersecurity risks so that potential breaches or attacks can be predicted ahead of time. Most security threats are programmed to operate based on history; hence, deep learning algorithms can analyze earlier occurrences and signs and then predict threats. For example, they can identify high risks involved with any individual, system, or activity similar to previous breaches so that precautionary measures can be undertaken. This has a predictive advantage and moves away from the reactive mode, saving time on the time it takes to respond to any innovative threats, hence preventing major disasters.
- **Automated Threat Classification and Incident Response:** Therefore, the timeliness with which an incident is detected and dealt with in the financial services sector is essential. Using deep learning algorithms, the identification of whether the occurred event could be categorized as a threat or not could be easily achieved. It helps SOCs steer their resources where they are most needed since criminals are not restricted to specific business sectors. By pre-filtering or pre-suspected categorization, deep learning models can eliminate insignificant alerts from sight and allow analysts to focus mainly on cyber incidents that are most likely to be critical and might need analysts' input. Also, it can cause deep learning in the ability to identify an incident response system that executes defined action whenever a breach has occurred, like a blockage of an IP address or isolation of vulnerable systems, and mitigate harm as it occurs.
- **Handling Complex and High-Dimensional Data:** There is an identification issue of high-dimensional and unstructured data, which becomes a challenge in traditional risk management methods. That is why deep learning, particularly the LSTM algorithm, is capable of handling such a complexity of the problem. For example, LSTMs can learn patterns in time-based sequences like transaction history to identify fraudulent or malicious activity. In addition, deep learning models can be used to filter and analyze data that may be in the form of text, images, or logs that are not easily manageable by conventional systems, which makes deep learning provide a better view of security threats.
- **Scalability and Adaptability:** Thus, risk management systems in cybersecurity shall also be scalable and flexible, designed to deal with the constantly increasing number of threats and their kinds. The scalability requirement is innate in most deep learning models and can handle large volumes of data and threats. They do not require manual intervention. Therefore, new strategies can be introduced easily, especially when analyzed from new data, which means they can adapt easily to new attack strategies. This dynamic nature is essential, especially for financial institutions, since the cyber environment threats are constantly changing; d hence, their organisations must always be on their guard.
- **Improved Accuracy and Reduced False Positives:** Another crucial problem of cyber risk management is an excessive number of false positives received from the existing systems that inevitably result in an analyst overload. Deep learning models, on the other hand, are better suited in these cases as they ensure a low number of false positives or, rather, an ability to accurately classify between malicious activities and normal activities. As these models adapt based on the information in a specific database, they become more effective in reducing the noise factor about true threats, thus presenting fewer but more accurate alerts. Since the risks associated with financial services are high, any excessive false alarms must be minimized so that the inputs can be well managed and the potential breaches can be addressed at the earliest opportunity.
- **Real-Time Risk Assessment and Mitigation:** Real-time threats are explicitly conspicuous for financial services, as a tiny delay in detecting them could be financially devastating. Deep learning models can easily process data faster than other models, responding to real-time threats. This is especially true in instances like checking for fraudulent activities in a company's transactions since this has to be done in real-time to stop such activities. Through real-time analysis, deep learning helps institutions to respond rapidly to threats, thus reducing any losses and negative impacts on their reputation.

1.2 Rise of Predictive Cyber Risk Management

Predictive cyber risk management is a revolutionary approach to managing cyber risks within an organization. For a long time, cybersecurity risk management was more or less just a branch of solving such crises as soon as possible once they happened. [4,5] Nevertheless, in the light of predictive risk management influenced by the use of Machine Learning (ML) and Deep Learning (DL), prospects of threats can be predicted, which allows organizations to avoid possible risks. Machine learning techniques also forecast events by looking into historical and real-time data for signs of a breach before it expands. Deep learning, a subset of machine learning, plays a crucial role in this transformation. This capacity to analyze intricate structures and abnormalities in limitless, mass, unorganized, and non-numerical data makes it suitable with regard to cyber risk prediction, including system log data, network data, and user data. Unlike conventional systems, where the features and the rules of learning are predefined, they adopt some form of learning from the data without external interference from the programmer.

Most importantly, with the ability to analyze huge amounts of data, deep learning mechanisms can find some correlation patterns that may be unnoticed by the analyst or hidden from initial observation by classical statistical tools, enhancing the accuracy of threat predictions. These models can recognize anomalous actions or behavior of a user, atypical usage of gaining system access, or other networking activities, which are also indicators of a forewarning of security threats. Thus, deep learning contributes effectively to moving organizations from a passive security stance where attacks and data breaches are anticipated and frequently occurring to an active security stance where such events are not expected and are significantly damaging if they occur at all. This means that predictive cyber risk management is beneficial in advancing the security of organizations and optimizing the flow of security to where most of the threats are.

2. Literature Survey

2.1 Overview of Cyber Risk in Financial Services

Cyber risk in financial services This is a dynamic threat given that the financial services segment has many valuable assets and customers' information is considered sensitive. These risks come from various sources: phishing attempts or scams, malware injection, ransomware attacks, internal threats, and DDoS assaults. [6-9] The financial industry, for example, is one of the most common targets and attacks that are likely to be affected by credential theft and social engineering. Generally, studies performed this year indicate that 80% of security breaches are due to stolen legitimizing information or social engineering fraudster tactics on the employees. This is due to the increasing complexity of the threats, which demands a more proactive and strong defense mechanism in financial organisations.

2.2 Traditional Risk Management Techniques

Traditional approaches and frameworks to the management of risks in cybersecurity include the NIST Risk Management Framework (RMF) and ISO/IEC 27005. These are systematic approaches since they offer a framework for establishing procedures for detecting, evaluating, and controlling cyber risks. Nevertheless, they rely chiefly on ratings, opinions of other professionals, and historical statistics. Although all such frameworks are good as they help maintain compliance and set up a strong risk posture, none of them provide real-time risk intelligence or predict breaches in advance. Another effect is that they are reactive and, hence, incapable of protecting against new cyber threats, given the dynamic, technologically advanced world.

2.3 Machine Learning in Cybersecurity

For the past two decades, Machine Learning (ML) has found its application in the modeling and classifying of intrusions, threats, and behavior in the field of cyber security. Experts have used different advanced supervised machine learning techniques such as SVM, decision trees, and random forests to identify anomalies or malicious activities in net traffic. Even though these models have exhibited a reasonable degree of accuracy, they encounter problems when working with more complex, noisy, or unstructured data sets that are common in cybersecurity. For instance, SVM is incapable of handling log data, which is unstructured text; hence, k-Nearest Neighbors (k-NN) models have high computational complexity and will take time to scale. Nevertheless, ML has made progress in automating threat identification and minimizing reaction time.

2.4 Deep Learning Applications

Recently, deep learning has emerged as an effective tool for the cybersecurity area, as it can automatically extract features from large datasets. Long Short-Term Memory (LSTM), a Recurrent Neural Network type, is known to effectively detect abnormality in sequential data such as system logs and activity logs. Likewise, Convolutional Neural Networks (CNNs) have also been applied for analysing the binary files and detecting the presence of malicious code patterns with good accuracy. These models are very good at modeling events or phenomena with complex dependencies but require no additional efforts to extract features from the given inputs. Nevertheless, although identifying reliable deep learning techniques to crack individual tasks is already possible, the incorporation into intensive breach predicting solutions remains restricted. The existing studies are insufficient and do not offer integrated methodologies incorporating different DL approaches for financial systems' preventive risk assessment and prediction.

3. Methodology

3.1 Data Collection

Collecting data for this study involved compiling data concerning cybersecurity from the most authoritative sources to compile a dataset that could be used for breach analysis. A third portion of the data was obtained from other sources, which are widely-known breach databases, such as the 'Verizon Data Breach Investigations Report (DBIR) 2019'. The DBIR offers practical information about some of the real security breaches in various industries, together with information about the techniques used by the attackers, victims, assets targeted, and timeframes. [10-14] This report significantly identified basic patterns and shared similarities of breaches within financial services. Besides the former, logs were collected from the IT department of cooperating selected financial organizations, and other data sources were anonymized. Such logs included system, authentication, firewalls, email gateway, and user trails logs.

This not only improved efficiency but also implemented the extra features of internal logs, which include several logs before an incident and the behavior before the incident happened. In order to eliminate the risks associated with the utilization

of identifiable data, the data was scrubbed and processed to replace any area that contains personally identifiable information (PII) in line with the GDPR and HIPAA guidelines. This was followed by temporal alignment and normalization of log data in order to have the same data format and time across different systems and time zones. In addition, the external threat intelligence feeds in the form of STIX/TAXII were incorporated into the dataset. These feeds provided such specific data as known threat actors, malware, C2, and IoCs that are critical in providing meaning to the logs collected and recognizing anomalous behaviors. This way, threat intelligence integration was more effective in labeling events for supervised learning. Combined with the fact that the collection of research sources was not limited to English-language publications, it can be stated that the sources collected provided good coverage of the topic regarding historical evolution and new threats in financial services.

3.2 Data Preprocessing

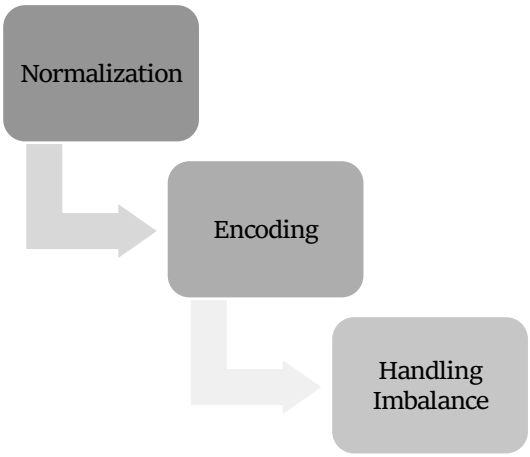

Fig 2: Data Preprocessing

- **Normalization:** In pre-processing the data, to try and stabilize the system during training and because there are significant differences in ranges of variables, additional data normalization was done using a min-max scaler. This technique standardizes the values between 0 and 1 by transforming them based on the minimum and maximum values for the given feature. This step avoids features with a wide range of numbers that overpower other features in the model and improves the efficiency of distance-based functions and neural networks.
- **Encoding:** For feature data such as the event type, users, and categorical regions, the data was encoded using the one-hot method. This method transforms a categorical value into a numerical one by making each category a feature column with its unique value. One-hot encoding is used because it differs from the preceding approach by not introducing additional order where it may not exist, which may pose a problem to algorithms that are specific with the numerical inputs.
- **Handling Imbalance:** Since the collections of cybersecurity data are imbalanced with many benign samples and a small number of real threats, the imbalanced problem was solved using the SMOTE method. As opposed to repeating the same samples, SMOTE synthesises samples of the minority class by creating a new one using linear interpolation between existent samples. It aids in augmenting model sensitivity to minimal but important breach incidents, which is the optimum feature in random samples that will help identify true positive cases in asymmetric distribution conduct.

3.3 Feature Engineering

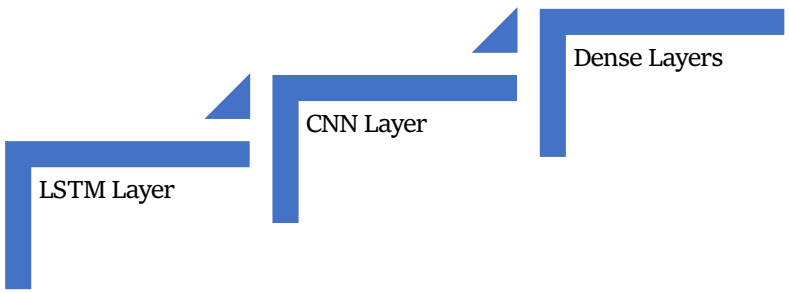

Fig 3: Feature Engineering

Feature engineering applies the right transformations and data manipulations to cybersecurity data to obtain inputs that can be fed to predictive models. This centred on the selection and formation of the features that would describe the users and the activity of the system, as well as the integration of threat intelligence. User activity parameters were also incorporated into the primary list of features. They comprised such values as login time and frequency, access to certain materials and objects, and non-routine activity of the user. For instance, logging into a system when not at work or from a remote locale can be a red flag, as can the application of wrong credentials or internal presumed threats. These behavioral indicators were accumulated in the time segments, which allowed for finding precedents to the security breach. Another important set of features was obtained from the activity log; this record depicted the running status of devices and servers in the system. Some of the observable parameters include the level of CPU utilization, the amount of memory used by a process, how frequently a file is accessed, and the frequency at which a process is executed, among others, to detect the characteristics of a process that may be involved in anomalous behaviors.

These surges or numerous accesses could clearly indicate the system containing malware, attempts at data extraction, or scanning that is not permitted. Having system-level characteristics assisted the model in gaining an architectural view of the specified threat profiles. Besides the behavioral and systems characteristics, threat intelligence feeds were also incorporated to give more context about the identified malicious actors. So, the Domain-listing IP addresses from blacklists, signatures of known malware families, and domains associated with known phishing initiations were correlated with internal logs. This enrichment process was truly beneficial in labelling interactions with the external threat actors and led to better predictions of them. In order to transform the threat indicators into the event type within the dataset, STIX/TAXII feeds were used, and the threat information was encoded either in binary or on a frequency basis. Altogether, these multi-dimensional features supplied a solid ground for assessing various feature-based signs pointing to possible internal breaches and external threats for the predictive models.

3.4 Model Architecture

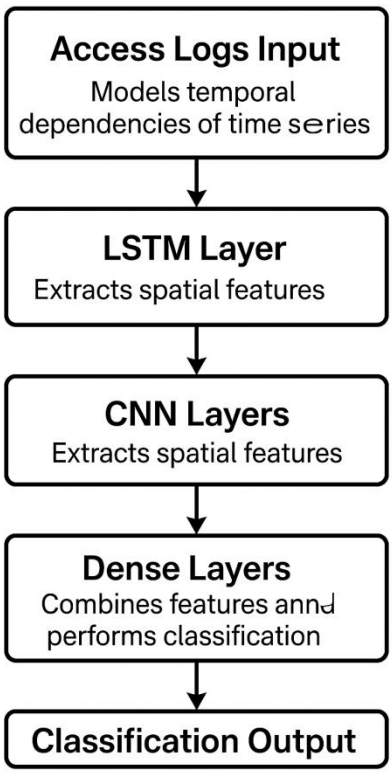

Fig 3: Model Architecture

- **LSTM Layer: The** LSTM network was adopted to model the temporal dependencies of time series to extract temporal relations from the access logs. LSTMs prove useful for mapping temporal sequences of the user and system events because, as has been found, people's behavior and the events happening in the system are temporal. Due to this kind of memory, LSTM can identify such scenarios as gradual shifts in behaviors, unusual login patterns, or a set of actions that are usually performed before a particular security threat. [15-17] This attribute proves vital in analyzing APTs and other insider attacks that occur for a long duration at a given online business.
- **CNN Layer:** Rot by 180° Deep Neural Networks such as CNNs were used for feature extraction of high-level abstract representations of the documents, particularly systems logs as documented in structured and semi-structured

documents. CNNs can detect spatial features or hierarchies in data, which makes them appropriate for detecting logarithmic sequences, bit strings, or frequency maps of hacker activities. The information on local interactions is drawn in by applying multiple filters and performing several pooling operations followed by the CNN layer, which is capable of retrieving such features as repeated unauthorized access attempts, as well as multiple uses of specific malware signatures, while conventional approaches may fail to do that.

- **Dense Layers:** The final stage of the architecture was the layers that acted as dense layers to combine the features that have been learned by the LSTM and CNN layers. These layers provided non-linear transformations to integrate temporal and spatial information and then fifteen features for classification. Using activations of softmax or sigmoid, depending on the type of the problem binary or multiple classes in the output layer, delivered probability values for each class label. These layers allowed the model to make effective probabilistic decisions about the possibility of experiencing a security breach, which made it the decision-making layer of the model's architecture.

3.5 Evaluation Metrics

For evaluation purposes of the developed breach prediction model, several parameters were used to give the model a comprehensive evaluation. Accuracy, sensitivity, specificity, and F1 score were the guidelines used while modeling to determine the model's capability in terms of true positive rate, false positive rate, and false negative rate. Recall, on the other hand, measures the ratio of actual breaches predicted to be breached out of all actual breached ones and enables determining if the model used provides broad and accurate results. Recall, on the other hand, takes the percentage positivity of true positives out of all actual positives; it reflects the ability of the model to identify as many breach events as possible. The F1 index is the arithmetic mean of the precision and recall metrics, as it considers the optimal ratio of true positive hits and false values. This is particularly relevant to cybersecurity because high recall (identifying as many threats as possible) without a steep amount of false positives (low precision) is valuable. Moreover, the Discrimination performance of the model was assessed using the ROC AUC measure. The ROC curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity) across various thresholds.

This Area Under the Curve (AUC) measures the model's actual performance in discriminating between positive and negative classes or instances, where the model with a value nearer to 1 has better accuracy. AUC stands out when the classes in the data set are disproportionate since it allows a proper measure of the model regardless of the chosen point for the classification. A final analysis was also made using the confusion matrix to check the model's performance based on the evaluation criteria where each class yielded true positives, true negatives, false positives, and false negatives count. This matrix gives clear information about the areas where the model is faltering. Why is the model making such misclassifications? Is it because the wrong classification is made, more breaches are non-breaches, or vice versa? Taking these metrics all together, one can have a more or less fair assessment of the model's predictive performance and the practical applicability of the approach to cybersecurity.

4. Results and Discussion

4.1 Model Performance

As mentioned earlier, accuracy is not the appropriate evaluation metric for imbalanced datasets; thus, four accuracy measurements were used, namely precision, recall, F1-Score, and ROC-AUC, using the three aforementioned models. These measurements clearly show how effectively each model works to identify cybersecurity breaches without making the most wrong decisions on such events.

Table 1: Performance Comparison

Model	Precision	Recall	F1-Score	ROC-AUC
SVM	74.2%	70.5%	72.3%	78%
Random Forest	78.5%	76.2%	77.3%	81%
LSTM-CNN (Proposed)	85.6%	84.7%	85.1%	91%

- **Precision:** Sensitivity measures the ratio of true breaches to the number of cases predicted as breaches (true breaches and false alarms). SVM had an accuracy of 74.2 %, meaning that at any point in time, the model deemed a system was breached, as it had been. This increased it slightly to 78.5% in the Random Forest model, which is a better feature for classifying malicious events. The proposed LSTM-CNN model achieved 85.6%, which was higher than both models and adorned more accuracy in discriminating between false alarms and genuine break-ins.
- **Recall:** Recall focuses more on measuring the model to accurately point out all the actual positive instances (actual breaches). The above SVM model obtained approximately 70.5% recall, meaning that out of all the actual breaches, it only pointed out about 70.5% while missing the others. Random Forest yielded a better result with a recall of 76.2% and, therefore, identified more breaches than the other models. In terms of recall, LSTM-CNN gained 84.7%; this means that most of the breaches were detected with a couple of false alarms being given.

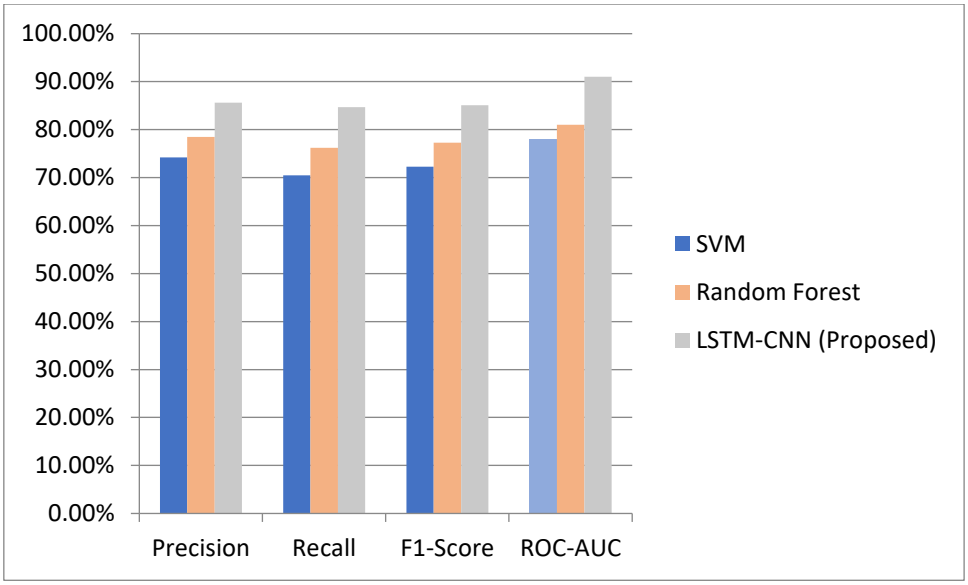

Fig 4: Graph representing Performance Comparison

- **F1-Score:** F1-Score is thereby defined as the proportion of the records that are relevant and classified as such by the model and, at the same time, represents the harmonic mean between precision and recall that allows balancing the model's precision/recall equation. The next model, SVM, had slightly lower precision and recall; therefore, the implementation of the SVM model had a moderate strike-section value of 72.3%. – The Random Forest kind bizarrely did comparatively better, with the models attaining 77.3% scores. The LSTM-CNN model had the highest F1-Score of 85.1% as it proved that it is capable not only of a high precision but also a high recall.
- **ROC-AUC:** The ROC-AUC measures the capacity of the model to correctly identify the positive class over the negative class. A higher value of AUC nearer to 1 signifies more accurate performance. The accuracy of the SVM model was 78%, which could be better in differentiating between breach and non-breach events. The results obtained for the Random Forest model were the AUC of 81%, which is considered an improved performance with higher discriminatory power. Thus, the LSTM-CNN hybrid model achieved the ROC-AUC of 91%, confirming its high performance in differentiating between security breaches and normal traffic.

4.2 Discussion

Of the reasons that have contributed to the high performance of the LSTM-CNN hybrid model, without a doubt, the biggest one is its ability to have a two-layered structure that consists of LSTM networks and CNNs. One of the primary uses of LSTM is its capability to model temporal dependencies, which are crucial characteristics of the cybersecurity event sequences. For example, login time and frequency remain constant; interaction with the system occurs over time. Hence, the temporal patterns have to be detected based on some change. Due to the capability of LSTM to retain information over the sequence, the model learns well the specific sequences that may likely point to a breach, such as log-ins at odd hours or repeated non-standard access attempts. The CNN layer is used to identify spatial features from structured logs such as access IP tally, system logs, and other types of structured event log data. CNNs are ideal for detecting any sequence or pattern/feature, such as multiple requests originating from an IP address, which may be an attacker, or observing various systems' behaviors of a familiar attack type, such as port scanning or credential stuffing.

This means that this is the only way for the model to detect very high-level patterns that isolate techniques that might be incapable of distilling from the immense and rather unorganized datasets. Furthermore, with a view of testing the stability of the model, the K fold cross-validation was performed on the dataset, where in this process, the entire database was split into many K subsets, and results were obtained K times each time, excluding one subset from the data. This is a violation of the hold-out method because it is a strategy to check the model validity on different new unseen data in every fold, thereby ascertaining the model's proficiency in diverse conditions and other distributions. Hence, using the LSTM-CNN hybrid model is not only very efficient in identifying breaches, but it is also highly dependable in guaranteeing that it can be utilized in real-world cybersecurity systems, which need accuracy, especially generalization ability.

4.3 Real-World Implications

- **Improved Incident Response Times:** Applying the LSTM-CNN combined model to the SOC can improve the response times to incidents without sacrificing the ability to detect new or falsified attacks. Therefore, it is accurate with high precision and recall and can take a relatively short time to detect potential security threats to allow security teams to respond appropriately. Minimizing false positives allows the alert provided by the analysts to not have an

immense amount to discuss with topics that are irrelevant to threats that are real. This means that there is less time between detecting a critical event and responding to it, which is useful in indicating the extent of the damage caused by the security violations in the system.

- **Dynamic Risk Scoring of Users and Endpoints:** However, the LSTM-CNN model has unprecedented opportunities to control an ongoing user's behavior and system interactions, which is the model's main advantage for dynamic risk score evaluation. However, by analyzing patterns such as frequent login times, abnormal use of data, or system interface, the model can generate real-time risk scores for the user and/or endpoint. Such scoring systems can also elicit certain actions, such as flagging a user with multiple threats, such as high risk, locking down access to the account, or blocking access to certain data if the behavior observed is suspicious. Dynamic risk rating enables an organization to adapt quickly to new threats given different risk levels and enhance the scores of access control measures.
- **Compliance Support:** For instance, the LSTM-CNN hybrid model can be of great use to organizations within industries that fall under GDPR and PCI-DSS regulations on data protection. In terms of regulation, the model is useful since it can promptly identify violations of security measures by providing continuous and automatic system monitoring. Moreover, the solution produces extensive logs and audit trails of security activities that benefit compliance assessment. He also said that these reports could be utilized to show compliance with data protection, ensuring that an organization has done all it could to protect customer data and diminish the chances of getting into trouble with the law.

5. Conclusion

This paper proposed an original context-aware deep learning model for detecting and preventing cyber data breaches in the financial services industry and a topic that is becoming relevant all the more as cybersecurity threats are becoming more frequent and advanced. The harsh fact is that in this study, the structure proposed utilizes LSTM and CNN to have a holistic view of the data since the LSTM component is responsible for temporal structures while the CNN component detects spatial structures. This two-tier approach enhanced the ability to use the model in a much better way to detect anomalous behaviour and make better forecasts compared to the standard method of machine learning algorithms like Support Vector Machines and Random Forests. It is evident that the proposed model achieved a notable improvement in the precision, recall, F1 score, and ROC AUC; all these measures are essential in determining whether or not the system can correctly detect security breaches while minimizing false positives and false negatives.

One is the practical implementation of the model, which had a successful result due to data integration of various data sources. These were the public breach datasets, internal IT logs, and external feeds; the model had a broad view of security threats. It also incorporated proper feature extraction, whereby other behavioural and activity indices related to the system were developed and used to increase the chances of breach prediction. Thus, widening the spectrum of analyzed data and encompassing various patterns allowed for the recognition of numerous interconnections between the various types of cybersecurity incidents, which enhanced the accuracy of the model.

In future work, the present work will be extended by deploying this framework in real-time, which is critical for the application in the field. Real-time monitoring and detection of security breaches help an organization contain the effects of cyber threats. If the model is to be used to process the data streams in near real-time, the current stream processing platforms like Apache Kafka could be used for this purpose to offer near real-time breach detection and response. This would allow security personnel to respond to threats in real-time so engagement time is eliminated, thereby reducing the breach's impact. Moreover, future work could focus on exploring the applicability of the introduced architecture to domains other than the financial services industry and specificity typical for different sectors and data kinds. In conclusion, this proposed research provides a good background that shows how deep learning models can improve cybersecurity architectures to quickly and effectively identify breaches in shifting situations.

References

[1] Ross, R. S. (2018). Risk management framework for information systems and organizations: A system life cycle approach for security and privacy.

[2] Sommer, R., & Paxson, V. (2010, May). Outside the closed world: On using machine learning for network intrusion detection. In 2010 IEEE symposium on security and privacy (pp. 305-316). IEEE.

[3] Buczak, A. L., & Guven, E. (2015). A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications surveys & tutorials, 18(2), 1153-1176.

[4] Anderson, R., & Moore, T. (2006). The economics of information security. science, 314(5799), 610-613.

[5] Bouyon, S., & Krause, S. (2018). Cybersecurity in Finance: Getting the policy mix right. Rowman & Littlefield.

[6] Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: techniques, datasets, and challenges. Cybersecurity, 2(1), 1-22.

[7] Sangkatsanee, P., Wattanapongsakorn, N., & Charnsripinyo, C. (2011). Practical real-time intrusion detection using machine learning approaches. Computer Communications, 34(18), 2227-2235.

[8] Zhang, J., & Zulkernine, M. (2006, June). Anomaly-based network intrusion detection with unsupervised outlier detection. In 2006 IEEE International Conference on Communications (Vol. 5, pp. 2388-2393). IEEE.
[9] Kim, G., Lee, S., & Kim, S. (2014). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Systems with Applications, 41(4), 1690-1700.
[10] Roy, S., Ellis, C., Shiva, S., Dasgupta, D., Shandilya, V., & Wu, Q. (2010, January). A survey of game theory as applied to network security. In 2010 43rd Hawaii International Conference on System Sciences (pp. 1-10). IEEE.
[11] Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. Ieee Access, 5, 21954-21961.
[12] Saxe, J., & Berlin, K. (2015, October). Deep neural network-based malware detection using two-dimensional binary program features. In 2015, the 10th International Conference on Malicious and Unwanted Software (MALWARE) (pp. 11-20). IEEE.
[13] De Gusmão, A. P. H., Silva, M. M., Poleto, T., e Silva, L. C., & Costa, A. P. C. S. (2018). Cybersecurity risk analysis model using fault tree analysis and fuzzy decision theory. International Journal of Information Management, 43, 248-260.
[14] Subroto, A., & Apriyana, A. (2019). Cyber risk prediction through social media big data analytics and statistical machine learning. Journal of Big Data, 6(1), 50.
[15] Mashrur, A., Luo, W., Zaidi, N. A., & Robles-Kelly, A. (2020). Machine learning for financial risk management: a survey. Ieee Access, 8, 203203-203223.
[16] Leo, M., Sharma, S., & Maddulety, K. (2019). Machine learning in banking risk management: A literature review. Risks, 7(1), 29.
[17] Ma, X., & Lv, S. (2019). Financial credit risk prediction in internet finance driven by machine learning. Neural Computing and Applications, 31(12), 8359-8367.
[18] Taplin, R. (Ed.). (2016). Managing Cyber Risk in the Financial Sector. Routledge, Taylor & Francis Group.
[19] Handa, A., Sharma, A., & Shukla, S. K. (2019). Machine learning in cybersecurity: A review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 9(4), e1306.