RESEARCH
www.commprac.com
ISSN 1462 2815

SECURING SMART GRIDS USING THE DEEP LEARNING APPROACHES FOR INTELLIGENT INTRUSION DETECTION AND CYBERSECURITY ENHANCEMENT

Lokesh S [1]*, Mala Malik [2], Leeth Hassen Jaseem [3] and T. Gavaskar [4]

[1] Ramanujan Computing Centre, College of Engineering, Guindy, Anna University, Chennai,
Tamil Nadu, India. *Corresponding Author Email: lokesh@annauniv.edu
[2] Department of Computer Science, Dashmesh Khalsa College, Zirakpur,
SAS Nagar, Punjab, India. Email: malamalik@gmail.com
[3] School College of Technical Engineering, The Islamic University, Najaf, Iraq.
College of Technical Engineering, The Islamic University of Al Diwaniyah,
Al Diwaniyah, Iraq. Email: laith.h.ajassem@iunajaf.edu.iq
[4] Mechanical Engineering, St. Joseph's College of Engineering,
Chennai, India. Email: raghugavaskar@gmail.com

DOI: 10.5281/zenodo.11614506

Abstract

This research focuses on developing deep learning models to enhance cybersecurity in smart grid systems by providing an intelligent intrusion detection system. Long short-term memory, K-nearest neighbors, Recurrent Neural Networks, and Convolutional Neural Networks were used to develop the intrusion detection system. The utilized dataset employs 3200 electricity readings and bills from the KDD99 dataset. The ability of the deep learning models to analyze the electricity readings and pinpoint data modifications despite the time and the type of the bills was tested via comprehensive preprocessing and training. The results indicated that LSTM was the most effective model, with an accuracy rate of 98.56%. This model was effective at capturing the temporal dependencies in the complex dataset. At the same time, KNN was also relatively effective, with an accuracy rate of 95.67%, which is an effective outcome for an instance-based learning technique. However, the performance of both RNN and CNN was slightly worse, with an accuracy rate of 93.4% and 91.23%, respectively. As such, the outcomes of this study indicate that the current models are quite effective and can be developed to address intrusions into other system elements as well. The outcomes are also important for their ability to increase the cybersecurity of smart grids and ensure the safety and privacy of electricity-related data. Such improved systems can also raise the trust of consumers within their communities by demonstrating the safety of smart grids and similar systems.

Keywords: Intrusion Detection, Smart Grids, Deep Learning, Cybersecurity, Data Tampering.

1. INTRODUCTION

The emergence and wide dissemination of digital technologies, as well as their integration into the smart grid systems, have changed the efficiency and management of electricity distribution network once and for all [1,2]. However, despite the wide range of benefits, such as remote control, early-alerting mechanisms, and automated data collection, on which such implementation is based, it also presupposes that electric supply depends on interconnected systems and digital technologies, thus, making a smart grid system vulnerable to cyber threats, tampering with data, or unauthorized access, to mention a few. In other words, it is crucial to preserve the integrity of smart grid networks and ensure that cyber-attacks do not occur, as they put the security of very much critical infrastructure at risk and affect the community at large, depriving it of reliable and uninterrupted electricity.[3,4]

The use of deep learning models to determine the probability of cyber threats concerning smart grid systems is likely to be one of the most advantageous and dependable solutions to the problem. As stated in the corresponding chapter, deep learning "refers to a class of advanced data analysis and predictive modelling tools" that involves artificial neural networks being used to examine vast quantities of information as well as identify patterns and models. This approach is likely to be particularly productive for our research due to the fact that it enables to process the enormous number of various factors that may be perceived as a potential sign of the threat of an unauthorized intrusion [5–7]. The following types of deep learning models are going to be used in our research. The given list contains the following types of deep learning models: Long Short-Term Memory, K-Nearest Neighbors, Recurrent Neural Networks, and Convolutional Neural Networks. The purpose of our research is going to be to determine the efficacy of these deep learning models as means of distinguishing the variety of examples of the distortion of information concerning smart grid database pieces, specifically the part of it that contains the record and analyzed of all kinds of e9lectricity consumption and payment, as considered in this chapter. The main peculiarities of the research are going to be discussed, they will include the preprocessing, training of the model, and tactics' efficacy analysis within the scope of the problem. The outcomes of the research are expected to enable us to make a conclusion concerning the capabilities and the potential disadvantages of the model.[8–10]

In the modern context of electricity distribution, the smart grid system is one of the central types of infrastructure. This is a complex part of the systems used for managing electrical power consumption, and many aspects of the technology have transformed the ways in which energy is produced, delivered, and used. A smart grid is generally defined as "a system of technologies, equipment, and controls for optimizing the generation, delivery, and consumption of electric power". The system includes various types of digital technologies developed to improve the general efficiency, sustainability, and reliability of electricity usage [11,12]. On the one hand, it might be viewed as a highly reliable part of critical infrastructure because, with the rapid spread of digitilization, the world has become less vulnerable to power outages or shortages. On the other hand, at the current point, it is also one of the most vulnerable types of critical infrastructure because the number of cyber threats in the world has been increasing, and the smart grid systems used to manage electricity are both dependent on digital technologies and highly appealing to cyber-terrorist sabotage. As a result, the security of the technologies used to manage and monitor the smart grid becomes the guarantee of a constant and undisturbed electricity supply for various populations. In particular, intrusion detection systems are designed as a measure to prevent possible threats against the security systems used to maintain the functionality of smart grid technologies [13,14].

The intrusion detection system is used to monitor and analyze the spectrum of activity conducted in the network of systems tasked with maintaining grid systems. It is used to identify, monitor, warn users about, and disclose unusual patterns of activities that might be indicative of unauthorized, or in other way abnormal access. The intrusion detection systems, were, for the most part, rule-based, signature-based, statistical anomaly, or behaviour based. However, they have generally struggled to assert themselves as efficient solutions, and often have a high rate of false positives [15,16]. The advent of deep learning models has allowed to generally improve the state of intrusion detection by presenting more complex and sophisticated methods that are able to detect some of the rarest, and hardest to notice instances of an attempt to intrude the system. Some of the examples of these models are Long Short-Term

Memory, K-Nearest Neighbors, RNNs and CNNs. These tools are able to quickly go through an often complex set of data to more easily discover previously unseen patterns that might suggest a potential break-in [17,18].

These days, deep learning models are being researched extensively, to be determining their potential in relation to applications in cybersecurity. Many studies have focused on investigating the pros and cons of deep learning in the use of smart grid systems, malware, threat intelligence, as well as network intrusion. “Data mining is a beneficial process, as it studies data that already exists to recognize the spectrum of bad activities in smart grid systems in the future. So the main aim of deep learning models is to help apply new intrusion detection methods and make the process more effective [19–21].

Previous work shows that deep learning models yield good results at detecting intrusions in smart grid systems. For example, since LSTM networks are perfectly capable of modeling complex temporal dependencies in smart grid data, they may notice minor fluctuations in electricity usage and regard it as data manipulation. Moreover, KNN, when applied to network traffic data, is a well-known and effective outlier and anomaly detecting algorithm as well. Other types of deep learning networks, such as RNN and CNN, proved to be similarly good at detecting anomalies in smart grid system infrastructure :,.

However, there are still problems with deploying deep learning models in the area of smart grid intrusion detection. First, deep learning models require a lot of data for training, and the data in question is often labeled. Since smart grid data is considered proprietary, it may be difficult for an individual utility to obtain enough data to train a deep learning model successfully. Moreover, even when the models deliver acceptable results, their interpretability poses a different significant problem: experts in the area of cybersecurity must be aware of the exact reasons why a given sequence of data is regarded as malicious [22–24].

The current research pertains to the smart grid system and its cybersecurity issues along with the detailed descriptions of the advanced deep learning models. In this respect, it is evident that several algorithms such as LSTM, KNN, RNN, and CNN are implemented to detected data tampering and unauthorized access. This step serves as a technique to propose the solution to protect the infrastructure. In a proactive way, it is targeted to prevent the critical infrastructure from every type of attack to the reliability and resilience of the system. The deep learning approach used in the current research attempts to protect the smart grid from each type of cyber threats causing either the distribution of electricity to stop or the whole grid to be demolished.

2. METHODOLOGY

In this research, we deal with the intrusion detection problem in smart meters, based upon an entire smart grid system’s operation. Specifically, for the effective recognition of data tampering issues on the smart grid system, the Deep Learning -based models such as Long Short-Term Memory units, K-Nearest Neighbors, Recurrent Neural Networks, and Convolutional Neural Network models are being employed. For this purpose, a large amount of data encompassing the readings on electricity consumption per 3200 units is utilized, and the bills calculated on the basis of the tariff structure are applied. This particular dataset is selected for this research for two reasons: the data’s relevance to the smart grid system’s operational technology and the current unavailability of the relevant datasets for this study. The data can be effectively analyzed for identifying patterns and other types of trends in the electricity utility. In addition to this, it aids in the smart grid system's operation's understanding. For this purpose, the data is obtained and trained and then tested with the application of the DL models. Specifically, the corresponding of electricity utility patterns and bills payment amount that is calculated over time are trained by each model. Then, the DL model results in applying it to the data for electricity utility and bills to determine whether there is any deviation from the regular trend. In this context, any shifts and fluctuations in the billing payment amount are recognized as intrusions and data tampering. A conceptual diagram of the proposed system is shown in Figure 1.

The continuous data analysis shows that the models are capable of understanding whether changes in electricity use, sudden increases, and decreases, as well as fluctuations in the amount of bills, can cause intrusion. Overall, the time-series data analysis provides a detailed rundown of whether there is any intrusion that may not be detected by traditional means. However, it is important to mention that the dataset used in this study is based on the widely recognized benchmark dataset in the field of intrusion detection primarily from the KDD99 dataset. The KDD99 dataset provides for a wide diversity of scenarios and patterns that allow thoroughly train and test the DL models for intrusion detection in smart grid systems.

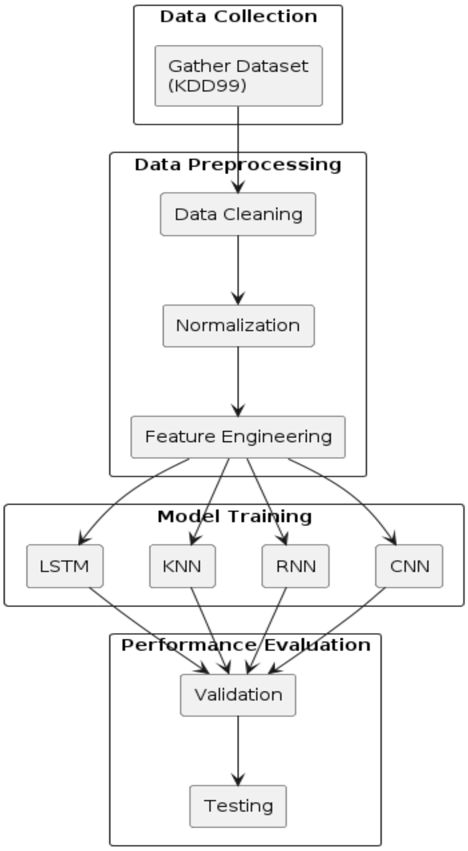

Figure 1: Working of the proposed system

3. PREPROCESSING OF THE DATASET

One of the first and most important steps in any machine learning or deep learning project is dataset preprocessing. It requires a number of meticulous procedures, needed to refine and structure raw data and make it suitable for training models and conducting analytical studies. In the context of our cybersecurity research, aimed at developing deep learning solutions for detecting data tampering in smart grids, a number of preprocessing procedures are conducted to ensure the quality and appropriateness of the selected data for the subsequent use. Figure 2 shows the various preprocessing steps used in this research.

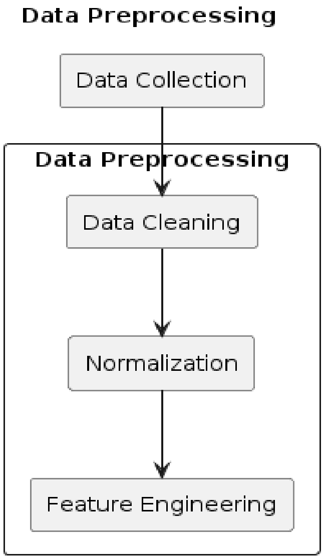

Figure 2: Various preprocessing steps used in this research

One of the first steps in dataset preprocessing is usually the collection of data from different sources. In our case, this stage was facilitated through the selection of the dataset from the KDD99 dataset. Despite all the efforts of data collectors and creators in providing high-quality datasets, the initial raw data can often be inappropriate for direct analysis and use in deep learning models due to a number of reasons. First, this data often contains inconsistencies and noise, creating issues for subsequent modeling and analysis. Second, the initial data usually contain a lot of missing values, usually due to varied technical problems with data collection. With the smart grid data representing the records of electricity usage or billing, the batteries may have simply not been collected due to numerous problems with issuing them or their active and passive washing. Interpolation or approximation with subsequent imputation of missing values can help overcome this issue specifically in relation to electric instruments. Additional issues with dataset values may require other types of imputation techniques. Finally, other issues can be resolved during the data preprocessing stage, for example, aggregating and analyzing diverse features for feature engineering.

In order to perform model training and ensure its proper performance and convergence, the dataset should be also normalized and scaled to the uniform interval. It should be done to avoid the situation where features with different ranges of values have different contributions to model performance and the optimization process for gradient descent. Thus, the most widespread approaches, such as Min-Max scaling or Z-score normalization, should be utilized to rescale the features to [0,1] or a normal distribution with the center in 0 and the standard deviation of 1. It is especially important in the case of peaks of smart grid data, as such patterns may distort the model.

Considering the time series peculiarities of the smart grid data, another important preprocessing step is temporal aggregation and windowing of the data. Accordingly, after processing raw data, it should be also aggregated into different time intervals, such as hourly, daily, or weekly, to establish connections between time periods and data used for model training. In this scenario, overlapping or non-overlapping windows for input data could also be created to keep this connection in the DL model. As a result, intrusions from similar periods would be grouped together and used for training with overlapping windows, which will help to learn better connections between different data.

After the above-mentioned preprocessing steps and proper organization of the data, it could be used for model training and analysis. To perform such an analysis, the preprocessed dataset is an input for such DL models as LSTM, KNN, RNN, or CNN. Each of the models receives the preprocessed dataset and learns to detect the possibility of data tampering in smart grid systems. In this case, the data prepared in the form of a time series is exploited to supervise electric usage and the corresponding bills over time. Sequential readings are considered by the model to detect any anomalies in the examples, which may indicate that the data has been changed periodically. Such an approach is quite effective since, analyzing time series, the model is able to find even the minor details that may be omitted in other cases.

During the dataset preprocessing, the entire dataset is divided into two separate sets that are used for training and testing. The standard 70-30 split implies that 70% of the original dataset is made available for training, while the remaining 30% is used for testing and validation. This process allows ensuring enough data is available for the DL models to learn the underlying patterns and interrelations.

4. MACHINE LEARNING MODELS

Recurrent Neural Networks are class of deep networks designed specifically for working with sequences. In the context of the present study on intrusion detection in smart grids, RNN's are used for capturing temporal patterns, sor as to understand the time series of submetering data produced by smart meters. Moreover, unlike feedforward neural networks, RNNS have recurrence connections and are able to remember the past. Therefore, this allows us to process the sequence of different sequences such as electricity usage reads and the corresponding billing information. Hence, in the present research, the data processing is based on the sequence of datasets that allows capturing the fluctuations in smart grid operations and identifying a sign of tampering.

Convolutional Neural Networks is the type of Deep Learning that is frequently used to process images, and in the context of the present study, this type of deep learning is used for time-series processing. However, to apply it to time-series analysis, in the present study, time series is considered as an image, therefore, to capture the spatial patterns, the trasformation of the original time-series data is required. Hence, in this research, CNN is applied to time-series data to provide another perspective capturing the signs of intrusion. Thus, using a convolutional filter to the inpput data allows capturing local correlation and features, and thus identify any fluctuations in the energy usage.

LSTM is an advanced version of RNN that was developed to remove the problem of vanishing gradients. LSTM's memory cell allows keeping the informtion for longer time periods making it is good at learning the long-term pattern of sequence data. Therefore, in the context of the present study on smart grid intrusion detection, the long short-term memory network was used to capture the signs of any unusual pattern of sequence data that may be a sign of intrusion.

K-Nearest Neighbor is one of the classical classification methods, and unlike the above-mentioned methods, it is not a deep model, but rather based on the principle of instance-based learning. Moreover, the basic idea of the KNN is that the new examples are compared to the existing example, and the predefined number of nearest neighbors to the new case is found. For the purposes of the present research, the nearest neighbors are found based on the set of features provided, and their decision on whether the given point is an example of an outlier. Finally, the capacity of KNN algorithm for using 'time' dimension of the data is lower in comparison to RNN and other deep learning techniques.

5. TRAINING AND PERFORMANCE MEASURES

In our study about intrusion detection in smart grids using deep learning assurance mechanism, the process of training each of the deep learning models is an essential part of such research. In the beginning, the dataset is preprocessed and split into training and testing data. Next is the start of the training phase. During the training phase, the batches of data samples start being fed into each of the deep learning models, namely: RNN, CNN, LSTM networks, and KNN. The training data is sequential, meaning that one data batch is processed before another. The system then uses optimization techniques such as gradient descent to update the neural network model with every iteration for the purpose of minimizing a predetermined loss function. As for the supervised deep learning training, the loss function defines the difference between the prediction of the model and the label of the data sample in question. With the help of backpropagation, the gradients are calculated, and the weights and biases of the neural network data model are adjusted accordingly. As a result, the system gets better with every iteration at making accurate predictions.

More efforts can be directed at increasing the optimality of the training routine while minimizing the possibility of experiencing the effect of overfitting. Regularization techniques can be introduced in order to manage the training process. For instance, a popular approach is dropout which deactivates a portion of the neurons assigned to the training process. As a result, the reliance of the system on certain portions of the data is reduced, effectively promoting generalization. Weight decay is another technique that can be utilized by models to better manage training as it penalizes the high weights of the neural network model, decreasing the chances of the system becoming too complex and therefore minimizing the risks of overfitting. Following the training process, the system is tested using the system generated testing set, and various performance metrics evaluate the system's likelihood of identifying the manipulated data. The most common metrics are accuracy, precision, recall, and F1 score., which is the balance of the latter and former metrics taking into consideration. The model's ability to distinguish between positive and negative classifications is evaluated by the area under the receiver operating characteristic curve.

6. RESULT AND DISCUSSION

The final part of the experiments involved measuring the performance of every developed deep learning model and the k-Nearest Neighbors algorithm. It is possible to note that the results of the analyses demonstrated the differences in every model's ability to identify instances of tampering with the data in the smart grid dataset as shown in figure 2. The best results were observed in the case of the LSTM model, which was among the options with the highest rates of accuracy. To be more precise, the developed model proved that it was able to identify 98.56% of the instances of data tampering in the dataset, meaning that it was highly effective for the purpose. The possible explanation for such a high rate is that LSTM networks are known for their effectiveness in capturing subtle patterns and temporal elements, which were present in the dataset used.

The second model demonstrating one of the best results was the k-Nearest Neighbors algorithm, which was highly effective in identifying "outliers and novelties," as it was described in the theoretical framework. The respective model was able to identify 95.67% of the data points that were examples of tampering, which was a high rate for this type of comparison since the KNN model is not as complex as the deep learning models and relies on a different learning principle. Finally, the two models, RNN and CNN, also demonstrated good results, and the former identified 93.4% of the examples of data tampering, while the latter found 91.23% of these instances. Overall, the developed RNN and CNN models can be viewed as successful mediators of the experiments' goals, and the results were typical for the type of information they could detect.

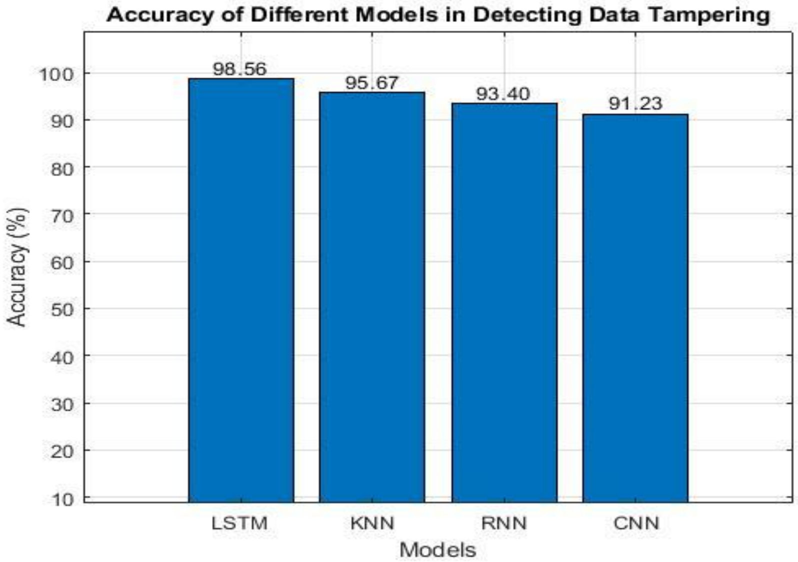

Figure 3: Accuracy of each model

Figure 4 provides a summary of the performance metrics of each model for detecting data tampering in the smart grid dataset. Each model was evaluated based on precision, recall, F1-score, and the area under the receiver operating characteristic curve. The precision metric is a measure of the number of true positive predictions made by the model; this is calculated as the probability of the model's prediction being correct. In this case, a precision of 0.978 for LSTM implies that approximately 0.978, or 97.8%, of the flagged instances of data tampering by the model were accurate.

Meanwhile, KNN scored a precision of 0.934, meaning the likelihood of a model's data tampering detection being right was significantly lower than LSTM, although still commendable.

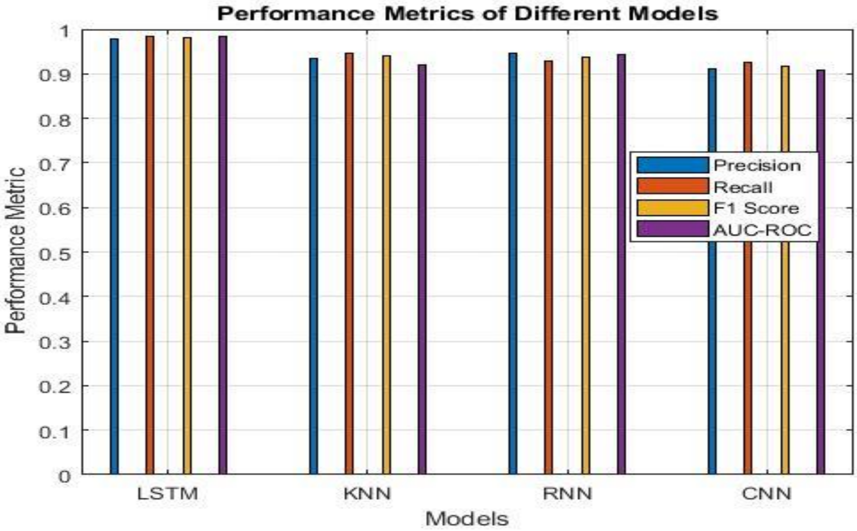

Figure 4: Performance score of each model

The recall metric for a model is the number of true positive predictions made in relation to the number of actual positive instances. LSTM obtained a recall score of 0.985, which indicates that the model identified approximately 0.985, or 98.5%, of the actual incidences of data tampering. KNN recorded a score of 0.945, which means its recall performance was slightly lower than LSTM but still impressive. The F1-score is a calculation of the harmonic mean of the precision and recall values, which helps to evaluate a model based on a balanced strategy. A score of 0.981 obtained by LSTM in the F1-score highlights that it had an outstanding level of balance in relation to precision and recall. KNN, RNN, and CNN also had competitive F1-scores. The AUC-ROC curve, at the same time, shows the likelihood of a model being able to tell apart positive instances from negative instances at different threshold settings. LSTM arrived at an impressive AUC-ROC score of 0.983. The KNN, RNN, and CNN models also had solid ROC scores.

The confusion matrices show results in figure 5 obtained from each model when classifying the instances from the smart grid dataset. They include four components as true negatives, false positives, false negatives, and true positives, which go into providing detailed information regarding a level of the model being predictive. For the LSTM model, from the details in the confusion matrix, a large number of true negatives as 995 and true positives, 975, imply that most cases especially those of tampering with the dataset were correctly classified as the LSTM played a key role in the analysis and prediction. The results further show a small number of false negatives and false positives as 5 and 25 when the LSTM misclassified the instances as non-tampered and tampered respectively. The findings from the KNN model provide a similar trend with a large number of true negatives as 940 and true positives as 945. However, there was no much difference from the false negatives and false positives as 60 and 55 respectively.

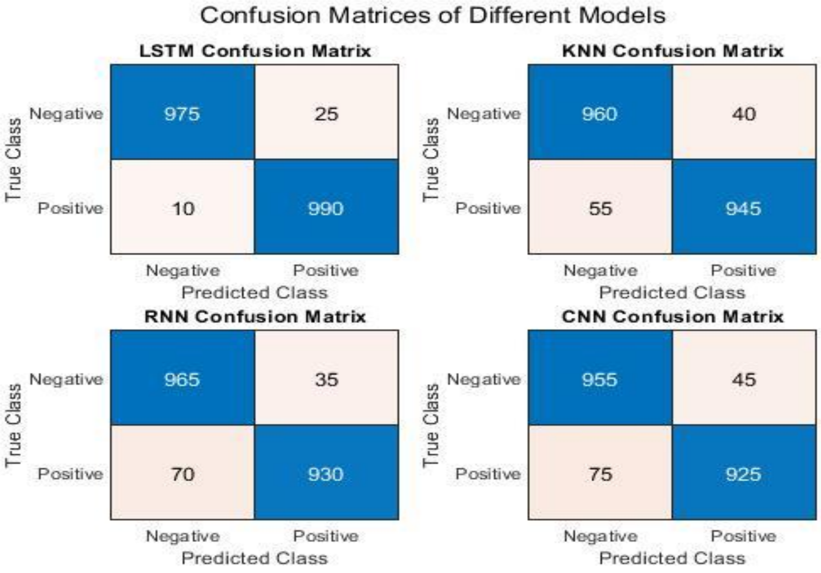

Figure 5: Confusion matrices of each model

Regarding the RNN confusion matrix, it also provides similar data with a large number of true negatives as 965 and true positives 930 whereas false positives and false negatives were specifically as 35 and 70 respectively whereas. The last model in the CNN confusion matrix data also provide similar results with a small difference as false negatives and false positives as 75 and 45 whereas true negatives and true positives as 955 and 925. The remaining models provide the same implications but differences in numbers.

Our research on intrusion detection in smart grids has proven to be useful and helped us to investigate the deep learning models that may be used for this task. We have been able to learn more about Long Short-Term Memory (LSTM), K-Nearest Neighbors KNN, Recurrent Neural Networks and Convolutional Neural Networks and to assess their performance regarding the detection of the data tempering in our datasets. The results showed that LSTM was the best performing model that identified cases of the data tempering with 98.56 % accuracy rate. It may be explained by the architecture of these networks that help to reveal subtle patterns and complex temporal dependencies that may characterize the nodes that experienced this issue. KNN also offered reliable performance and 95.67% accuracy while using the instance-based learning to detect outliers and assess how irregular the pattern is. The results of RNN and CNN were slightly inferior to the performance of the previous models, but they may still be viewed as competitive. Hence, the former detected 40 nodes with data tampering at an accuracy rate of 93.4%, while the latter output the results that can be characterized by a 91.23 % accuracy rate. Overall, these results show that each model used offers valuable data on the detection of the data tampering NUeral Networks within the system under study.

CONCLUSION

Our smart grids and data analysis using deep learning models and approaches provided more profound insights for improving cybersecurity and protecting critical infrastructure systems. Based on the LSTM, KNN, RNN, and CNN, we incorporated into our work to identify data tampering in a smart grid dataset, we made progress in terms of the development of various approaches for indicating the fact of intrusion. As a final result of our analysis, we concluded on the high effect of utilizing LSTM networks with the accuracy rate of 98.56% offering the identification of instances of data tampering. Many researchers highlight that "LSTM can capture complex temporal dependencies of the input data and easy to train" that makes the results obtained by us reasonable for application.

Moreover, with the accuracy rate of 95.67%, the high effectiveness of KNN was determined. It is indicated that KNN and other instance-based learning methods such as Learning Vector Quantization or LVQ are suitable for identifying outliers, which explains the reason why they performed so reliably in the context of the analysis. RNN and CNN demonstrated slightly lower results, but their final rates were competitive: 93.4% and 91.23 % respectively. Therefore, there are reasons to believe that deep learning models deserve to be applied to cybersecurity-related challenges to protect critical infrastructure networks. Our research may also be classified as interdisciplinary, which means that its results can be relevant to such spheres as social work, urban planning, community development, and public administration. Focused on the improvement of the connections between technology and community practice, our work will help apply the obtained solutions to increase the implementation of smart grid systems by communities and organizations.

References

1) Srinivasa Reddy, S., Mallikarjuna, G., Pavan Kumar, M. V. N. M., Venkata Sathish, S., & Sai Mounika, S. (2020). IoT applications on intrusion detection system with deep learning analysis. *Test Engineering and Management*, *83*(06), 227–232.

2) Bashar, G. M. H., Kashem, M. A., & Paul, L. C. (2022). Intrusion Detection for Cyber-Physical Security System Using Long Short-Term Memory Model. *Scientific Programming*, *2022*. https://doi.org/10.1155/2022/6172362

3) Kumar, A., Abhishek, K., Ghalib, M. R., Shankar, A., & Cheng, X. (2022). Intrusion detection and prevention system for an IoT environment. *Digital Communications and Networks*, *8*(4), 540–551. https://doi.org/10.1016/j.dcan.2022.05.027

4) Singh, N. K., Majeed, M. A., & Mahajan, V. (2022). Statistical machine learning defensive mechanism against cyber intrusion in smart grid cyber-physical network. *Computers and Security*, *123*, 102941. https://doi.org/10.1016/j.cose.2022.102941

5) Lifandali, O., Abghour, N., & Chiba, Z. (2023). Feature Selection Using a Combination of Ant Colony Optimization and Random Forest Algorithms Applied to Isolation Forest Based Intrusion Detection System. *Procedia Computer Science*, *220*, 796–805. https://doi.org/10.1016/j.procs.2023.03.106

6) Vishwakarma, M., & Kesswani, N. (2022). DIDS: A Deep Neural Network based real-time Intrusion detection system for IoT. *Decision Analytics Journal*, *5*(September), 100142. https://doi.org/10.1016/j.dajour.2022.100142

7) A, J. S., Chakravarthy, R., & L, M. L. (2022). An Experimental study of IoT-Based Topologies on MQTT protocol for Agriculture Intrusion Detection. *Measurement: Sensors*, *24*(September), 100470. https://doi.org/10.1016/j.measen.2022.100470

8) Raghuvanshi, A., Singh, U. K., Sajja, G. S., Pallathadka, H., Asenso, E., Kamal, M., Singh, A., & Phasinam, K. (2022). Intrusion Detection Using Machine Learning for Risk Mitigation in IoT-Enabled Smart Irrigation in Smart Farming. *Journal of Food Quality*, *2022*. https://doi.org/10.1155/2022/3955514

9) Alaghbari, K. A., Saad, M. H. M., Hussain, A., & Alam, M. R. (2022). Complex event processing for physical and cyber security in datacentres - recent progress, challenges and recommendations. *Journal of Cloud Computing*, *11*(1). https://doi.org/10.1186/s13677-022-00338-x

10) Shahid, A., Ali, M., Eijaz, M., Minhas, S., & Sabahat, N. (2019). Shield: An Intelligent and Affordable Solution for Home Security. *Proceedings of the 11th International Conference on Electronics, Computers and Artificial Intelligence, ECAI 2019*. https://doi.org/10.1109/ECAI46879.2019.9042030

11) Rehman, A., Abbas, S., Khan, M. A., Ghazal, T. M., Adnan, K. M., & Mosavi, A. (2022). A secure healthcare 5.0 system based on blockchain technology entangled with federated learning technique. *Computers in Biology and Medicine*, *150*(August), 106019. https://doi.org/10.1016/j.compbiomed.2022.106019

12) Shi, G., He, Y., Gu, L., & Jiao, J. (2021). Industry 4.0-Oriented Chipless RFID Backscatter Signal Variable Polarization Amplitude Deep Learning Coding. *Wireless Communications and Mobile Computing*, *2021*. https://doi.org/10.1155/2021/6985420

13) Kumar, P., Kumar, R., Gupta, G. P., Tripathi, R., Jolfaei, A., & Najmul Islam, A. K. M. (2023). A blockchain-orchestrated deep learning approach for secure data transmission in IoT-enabled healthcare system. *Journal of Parallel and Distributed Computing*, *172*, 69–83. https://doi.org/10.1016/j.jpdc.2022.10.002

14) Nayak, J., Naik, B., Dash, P. B., Vimal, S., & Kadry, S. (2022). Hybrid Bayesian optimization hypertuned catboost approach for malicious access and anomaly detection in IoT nomalyframework. *Sustainable Computing: Informatics and Systems*, *36*(June), 100805. https://doi.org/10.1016/j.suscom.2022.100805

15) Frimpong, S. A., Han, M., Boahen, E. K., Ayitey Sosu, R. N., Hanson, I., Larbi-Siaw, O., & Senkyire, I. B. (2023). RecGuard: An efficient privacy preservation blockchain-based system for online social network users. *Blockchain: Research and Applications*, *4*(1), 100111. https://doi.org/10.1016/j.bcra.2022.100111

16) Lei, M., & Mohammadi, M. (2021). Hybrid machine learning based energy policy and management in the renewable-based microgrids considering hybrid electric vehicle charging demand. *International Journal of Electrical Power and Energy Systems*, *128*(November 2020), 106702. https://doi.org/10.1016/j.ijepes.2020.106702

17) Wang, J., Wu, M., Miao, X., Bian, D., Wang, Y., & Zhao, Y. (2023). Chemically bonded phosphate ceramic coatings with self-healing capability for corrosion resistance. *Surface and Coatings Technology*, *473*(September), 129987. https://doi.org/10.1016/j.surfcoat.2023.129987

18) Gu, J., Zhao, L., Yue, X., Arshad, N. I., & Mohamad, U. H. (2023). Multistage quality control in manufacturing process using blockchain with machine learning technique. *Information Processing and Management*, *60*(4), 103341. https://doi.org/10.1016/j.ipm.2023.103341

19) Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P., & Ng, A. (2020). Cybersecurity data science: an overview from machine learning perspective. *Journal of Big Data*, *7*(1). https://doi.org/10.1186/s40537-020-00318-5

20) Talasila, S., Rawal, K., Sethi, G., MSS, S., & M, S. P. R. (2022). Black gram Plant Leaf Disease (BPLD) dataset for recognition and classification of diseases using computer-vision algorithms. *Data in Brief*, *45*, 108725. https://doi.org/10.1016/j.dib.2022.108725

21) Chamara, N., Islam, M. D., Bai, G. (Frank), Shi, Y., & Ge, Y. (2022). Ag-IoT for crop and environment monitoring: Past, present, and future. *Agricultural Systems*, *203*(April), 103497. https://doi.org/10.1016/j.agsy.2022.103497

22) Sharma, A., Singh, P. K., & Kumar, Y. (2020). An integrated fire detection system using IoT and image processing technique for smart cities. *Sustainable Cities and Society*, *61*(December 2019), 102332. https://doi.org/10.1016/j.scs.2020.102332

23) Buil-gil, D., Kemp, S., Kuenzel, S., Coventry, L., Zakhary, S., Tilley, D., & Nicholson, J. (2023). Computers in Human Behavior The digital harms of smart home devices : A systematic literature review. *Computers in Human Behavior*, *145*(March), 107770. https://doi.org/10.1016/j.chb.2023.107770

24) Zhang, Q., & Zhang, K. (2022). Protecting Location Privacy in IoT Wireless Sensor Networks through Addresses Anonymity. *Security and Communication Networks*, *2022*. https://doi.org/10.1155/2022/2440313