information

Article

Rapid Forecasting of Cyber Events Using Machine Learning-Enabled Features

Yussuf Ahmed [1,*], Muhammad Ajmal Azad [1] and Taufiq Asyhari [2]

[1] School Computing, Birmingham City University, SteamHouse, Belmont Row, Birmingham B4 7RQ, UK; muhammadajmal.azad@bcu.ac.uk
[2] Data Science Indonesia, Monash University, Green Office 9 Building, Jl. BSD Green Office Park, Sampora, Cisauk, Tangerang Regency, Banten 15345, Indonesia; taufiq.asyhari@monash.edu
* Correspondence: yussuf.ahmed@bcu.ac.uk

Abstract: In recent years, there has been a notable surge in both the complexity and volume of targeted cyber attacks, largely due to heightened vulnerabilities in widely adopted technologies. The Prediction and detection of early attacks are vital to mitigating potential risks from cyber attacks and network resilience. With the rapid increase of digital data and the increasing complexity of cyber attacks, big data has become a crucial tool for intrusion detection and forecasting. By leveraging the capabilities of unstructured big data, intrusion detection and forecasting systems can become more effective in detecting and preventing cyber attacks and anomalies. While some progress has been made on attack prediction, little attention has been given to forecasting cyber events based on time series and unstructured big data. In this research, we used the CSE-CIC-IDS2018 dataset, a comprehensive dataset containing several attacks on a realistic network. Then we used time-series forecasting techniques to construct time-series models with tuned parameters to assess the effectiveness of these techniques, which include Sequential Minimal Optimisation for regression (SMOreg), linear regression and Long Short-Term Memory (LSTM) to forecast the cyber events. We used machine learning algorithms such as Naive Bayes and random forest to evaluate the performance of the models. The best performance results of 90.4% were achieved with Support Vector Machine (SVM) and random forest. Additionally, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) metrics were used to evaluate forecasted event performance. SMOreg's forecasted events yielded the lowest MAE, while those from linear regression exhibited the lowest RMSE. This work is anticipated to contribute to effective cyber threat detection, aiming to reduce security breaches within critical infrastructure.

Keywords: forecasting; big data; time series; cyber attack prediction; cyber events; intrusion detection

Citation: Ahmed, Y.; Azad, M.A.; Asyhari, T. Rapid Forecasting of Cyber Events Using Machine Learning-Enabled Features. *Information* **2024**, *15*, 36. https://doi.org/10.3390/info15010036

Academic Editors: Jiaping Gui and Futai Zou

Received: 5 December 2023
Revised: 30 December 2023
Accepted: 5 January 2024
Published: 11 January 2024

1. Introduction

The threat landscape is dynamic and continuously evolving. It challenges even the best security defences deployed by organisations that invested a significant amount of their budget on security investments. Cybercriminals are finding ways to circumvent these security controls. The vast number of applications used in typical organisations also increases the attack surface due to potential vulnerabilities and the discovery of new software bugs. Security solution providers are also a target of these cyber attacks, as demonstrated by the attack that compromised a global cybersecurity firm, which was compromised through the SolarWinds update [1].

With the rapid and exponential growth of data generated by various sources such as network traffic logs, raw IP traffic, system logs, sensor data and Internet of Things (IoT) traffic, big data has become an important tool for the timely detection and identification of intrusions across heterogeneous networks. The marriage between big data, machine learning and artificial intelligence helps to collect, process and store a large volume of

unstructured data for real-time analysis and the extraction of meaningful information [2,3]. This meaningful information could be used to identify malicious traffic and block it effectively to minimise financial losses and reputational damage because of malicious traffic. Accurately predicting cyber attacks remains challenging due to the sophistication of the attacks and the large attack surface, which gives the cyber attacker many points of entry.

The trend is to move away from traditional security approaches and embrace predictive methods capable of detecting sophisticated attacks [4]. The cost of cyber attacks continues to increase, and it will likely follow the same trajectory. The average cost of a data breach is estimated to be approximately 4.35 million US Dollars (USD), and it takes around 277 days to detect and contain the breach [5]. There is a substantial increase in security investments to protect critical assets, but cybercriminals are finding ways to bypass these security defences.

The intensity of cyber attacks is likely to continue due to the vast number of connected devices, a large amount of unstructured data and heterogeneous network connectivity, which contribute to increased attack exposure. These devices range from tiny sensors to mobile devices capable of generating a massive amount of data that crosses the boundary of the network. Organised cyber criminals also use Advanced Persistent Threats (APT) to go beyond security defences. APT groups utilise complex and sophisticated techniques to avoid detection. However, several works exist to improve the detection capabilities of APT attacks [6–8]. Several authors and commercial providers have proposed predictive methods for detecting complex attacks and helping pre-empt such attacks. However, these are still works in progress, and accurate detection of complex attacks remains a challenge. The predictive approaches will help redirect technical resources to where they are needed. Such an effort will help to prevent data breaches and free valuable time for technical teams, allowing them to use their time more efficiently and deal with cyber incidents that require urgent attention.

This paper presents our work on cyber event forecasting to help accurately detect cyber attacks. Forecasting has been applied in other fields but is a developing area of research when it comes to applying it in the context of cyber attacks. This work contributes to these attempts to bring the use of unstructured big data along with rigorous ML approaches to cyber forecasting and improve accurate detection and prevention of cyber intrusions [9–11]. The proposed approach can use big data technologies to analyse network traffic logs in real time to identify malicious activities such as Denial of Service (DoS) attacks, malware infections and data exfiltration attempts. We also use machine learning algorithms along with big data to automatically classify the traffic into malicious and non-malicious traffic.

Our Contributions

Although forecasting is well established in other domains, such as weather and stock predictions, it is still an emerging area for cyber attack prediction [12]. Most of the existing work is based on social media feeds [13,14] and honeypots [15,16], which has its limitations given they often look at single attacks, such Denial of Service (Dos) and malware variant. Most of the work on forecasting is limited by the quality of the datasets. We used a large dataset captured from a realistic network to overcome these challenges. The dataset has multiple attack labels and is very comprehensive [17]. We performed data preparation and cyber event forecasting to predict cyber attacks within a specific time frame. The main contributions of our research are as follows:

- Perform time series resampling based on original big data to make sure we have an equal sample for forecasting and intrusions.
- Perform and evaluate time series forecasting based on linear regression, SMOreg and LSTM.
- Evaluate the performance of the forecasted events using the metrics MAE and RSME.
- Use time-series data to forecast cyber attack events within a specified period.

The rest of the paper is organised as follows. Section 2 discusses the related work in the area of intrusion detection, forecasting and the use of machine learning for classifying network traffic. Section 3 discusses the proposed approach for event forecasting and intrusion detection. Section 4 provides the experimental setup and dataset used for the evaluation. Section 5 presents the evaluation of the results, and Section 6 concludes the paper.

2. Related Work

Recent years have seen an increase in the number and volume of data breaches due to the availability of sophisticated tools and complex attacks from groups affiliated with state actors and organised criminals. The research community and industry have been working together to come up with solutions to address these security challenges, particularly in predicting cyber events more accurately. This paper contributes to that body of knowledge and aims to forecast cyber attacks based on certain cyber events and features on the network. We utilise data-driven approaches to predict these events before they occur to help the security teams better respond to such threats.

The next part will cover intrusion detection, including a detailed overview of the existing research.

2.1. Intrusion Detection

In recent times, the escalation of cyber attacks has prompted efforts aimed at identifying and preventing these intrusions with varying degrees of success. Diverse technologies, such as Intrusion Detection Systems (IDS), Intrusion Prevention Systems (IPS), Security Information and Event Management Systems (SIEMS), firewalls and anti-virus systems have been implemented to detect attacks and notify security teams. While these tools play a pivotal role in detecting and preventing cyber attacks, they are susceptible to generating false alerts, and accurately pinpointing sophisticated attacks remains a persistent challenge [18]. To combat cyber intrusions, several methodologies have emerged, primarily classified into two categories: signature-based intrusion detection systems and anomaly-based intrusion detection systems. Signature-based detection is effective against attacks with known signatures, while anomaly-based detection excels in identifying new attack patterns. Intrusion Detection Systems (IDS) are broadly categorised into three types: Network Intrusion Detection Systems (NIDS), Host Intrusion Detection Systems (HIDS) and Hybrid Intrusion Detection Systems. Among these, Network Intrusion Detection Systems (NIDS) represent the most widely embraced category of IDS, tasked with analysing network traffic to spot anomalies. Upon detection, these systems generate security alerts that are then prioritised and addressed by the security team. Examples of NIDS include Zeek [19] and Snort [20]. Researchers have explored the use of Machine Learning (ML) and Deep Learning (DL) methodologies to enhance the detection capabilities of NIDS. ML and DL-based NIDS models typically rely on datasets and usually encompass multiple stages, which are (i) data preparation, (ii) training and (iii) testing. In the data preparation stages, the dataset is prepared to make it suitable for machine learning, and it is then split into training and testing portions. Several authors have proposed NIDS models, but researchers are still working on improving the detection accuracy and minimising false alarms. In [21], the authors proposed a model based on deep learning approaches for network intrusion detection and utilised sparse auto-encoders. They trained the model to classify network traffic into benign and attack, but the approach was tested using binary classifications. In [22], the authors proposed a network intrusion detection model and utilised unsupervised autoencoders. They used a heuristics threshold to improve the detection accuracy of their proposed IDS. Reference [23] proposed an intrusion detection system using the Ensemble Core Vector Machine (CVM) approach to detect various types of attacks, including probe and DoS attacks. According to the authors, the model achieved a high accuracy result.

Host Intrusion Detection Systems (HIDS) detect anomalies in host systems and generate alerts. This is mainly installed on critical systems where security protection is essential.

It also helps collect indicators of compromise following suspicious activities reported by the HIDS system. Examples of such activities include unauthorised access attempts and unauthorised modification of files or programs. It is always good to correlate HIDs logs with other monitoring tools to help prioritise genuine threats. Examples of HIDS include Splunk [24] and Open Source Security Event Correlator (OSSEC) [25]. Several authors have carried out work improving the accuracy of HIDS. In [26], the authors proposed a HIDS model for cloud computing. The model alerts users when suspicious activities are detected based on systems called traces and classifies them using a KNN classifier. In [27], the authors proposed the HIDS model for Supervisory Control and Data Acquisition Systems (SCADA). Reference [28] used a combination of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) detection models, which led to an improved detection result.

Hybrid intrusion detection systems amalgamate two or more methods to enhance intrusion detection, diverging from conventional IDS approaches reliant on either signature-based or anomaly-based detection. Numerous researchers have introduced models in this domain. For instance, ref. [29] suggested a hybrid IDS model specifically designed to identify cyber attacks on the web. Their method combined signature-based and anomaly detection, achieving an accuracy rate of 96.7%. Similarly, ref. [30] proposed a model integrating anomaly-based and signature-based approaches to identify attacks on IoT networks. Their model encompassed three stages: traffic filtering, preprocessing and a hybrid IDS. In another instance, ref. [31] presented a hybrid IDS detection model for IoT, targeting the detection of Denial of Service (DoS) attacks and network traffic analysis. Any deviations from the standard were classified as potential attacks. Reference [32] proposed a hybrid architecture for IDS tailored for the Internet of Vehicles. Their architecture, based on Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), merged several datasets containing DDoS attacks and car hacking incidents to assess their model's performance. Their model achieved an overall detection accuracy of 99.5% and 99.9% for DDoS and car hacking, respectively. Lastly, in [33], the authors introduced a cyber kill chain-based hybrid IDS framework for a smart grid. They applied the cyber kill chain to identify cyber attacks at different stages of the chain.

While recent advances have seen an increase in the deployment of machine learning and deep learning approaches for improving detection accuracy, these models' accuracy depends on the quality of the datasets used. Some of the prominent IDS datasets include the KDD99 and NSL-KDD [34], which contained features that were used to differentiate normal and abnormal traffic. Other datasets that have been widely used include the Kyoto [35], UNSW-NB15 [36] and CICIDS-2017 [34] and CIC-IDS-2018 [34] datasets. Most of the work on intrusion detection research has been based on using machine learning data and using classification and performance metrics, such as percentage accuracy. For example, most of the work on the datasets above has used ML and DL approaches to extract features and perform feature engineering and classification to fine-tune the parameters to achieve the best accuracy results. Our work explored cyber event forecasting, which has not been explored widely in the cyber domain, and the forecasting work is not there to replace intrusion detection but to complement it.

Next, we will cover cyber event forecasting, predictions and related work. We will also briefly cover some of the other domains where forecasting has been applied and use it to inform our work.

2.2. Forecasting and Predictions

Researchers have recently shown interest in cyber attack forecasting, and their work is contributing to the body of knowledge. Their work ranges from survey papers to machine learning models and achieving varying results. Most of this work is on sentimental analysis and based on social media feeds, although others are looking at other attacks, such as DoS and malware variants. In [37], the authors performed cyber attack forecasting using machine learning techniques using data breaches spanning over 12 years. They analysed

the data and found the threats of cyber attacks to increase in frequency but not magnitude. Reference [38] used machine learning to predict the cost of cyber breaches with the view that their work could also be used to predict premiums in cyber insurance. References [39,40] used sentimental analysis to predict cyber attacks. In Ref. [41], the authors proposed a method aimed at aiding incident responders in predicting the possible functionalities of malware post-detection. Their methodology is grounded in a probabilistic model, empowering the forecast to recognize a range of capabilities and gauge the probability of each capability being executed. As per the authors, their approach not only unveils potential capabilities but also assigns weights based on the likelihood of their execution. Ref. [42] conducted an assessment of predictive methods' capabilities in the field of cybersecurity. Their proposed method aimed to identify potential attackers through the utilization of network entity reputation and scoring mechanisms. Ref. [43] provided an an overview of prediction and forecasting techniques employed in the realm of cybersecurity. Their focus was on the predicting the intention of the attackers and anticipating potential attacks that might impact the the overall security status of the network. Ref. [44] examined the present research directions concerning cyber attacks by scrutinizing the data-driven methodologies utilized by researchers in this swiftly evolving domain. Additionally, highlighted challenges and potential future trajectories within this field.

Time series-based techniques are widely adopted for forecasting future events. Such techniques are based on autoregressive time series and other models based on neural networks. Other well-known forecasting methods include ARIMA, linear regression, SMOreg, Gaussian process and multilayer perceptron. Reference [45] proposed time series-based anomaly techniques for dealing with adversarial attacks. Author [46] carried out a review of time series-based anomaly detection techniques and found there was no single technique that outperforms the others. Reference [47] applied time series techniques to build their predictive model. The model was used to detect vulnerabilities in internet browsers.

2.2.1. ARIMA

ARIMA, a statistical technique utilising time series data for future trend prediction, was explored in a study by [48]. The authors studied the data of the given parameters to improve the forecasting using ARIMA and Exponential Smoothing (ETS). The two forecasting methods were compared using parameters such as pressure and humidity. The accuracy was also compared using metrics such as MAE (Moving Absolute Error) and RMSE (Root Mean Square Error). ARIMA has been used for a long time, although there are some limitations with ARIMA models and, in particular, the difficulty of modelling nonlinear relationships [49]. Reference [50] used ARIMA-based forecasting to predict future cyber attacks based on historical incidents.

The authors [51] surveyed the prediction techniques used in cyber security and concluded their effectiveness is linked to the context in which they are used and the research direction. Reference [52] proposed a time series technique for predicting data breaches based on the size and incident time derived from historical data. They used Seasonal Autoregressive Integrated Moving Average (SARIMAX) and Recurrent Neural Networks (RNNs), and both models achieved good performance results. Reference [53] studied ARIMA and SARIMA models and evaluated them for long-term runoff forecasting. The results showed that the SARIMA model performed better than ARIMA at forecasting the annual runoff. ARIMA is a suitable statistical method for forecasting and only requires time series data, although the data has to be stationary.

2.2.2. Linear Regression and SMOreg

This algorithm predicts the correlation between two features and evaluates their connection [54]. Typically, there are dependent and independent variables involved. SMOreg offers an SVM-based solution for handling regression problems, excelling particularly in modelling and predicting with non-linear data [55].

2.2.3. Deep Learning

Deep learning (DL) is within the realm of machine learning and typically consists of multiple layers, including a hidden layer, which allows it to learn from the feature representations [56]. Several deep learning algorithms exist, including neural networks, Convolutional Neural Networks (CNNs), Long Short-Term Memory Networks (LSTMs) and Autoencoders.

Several authors have proposed models for detecting cyber attacks based on CNN. For example, ref. [57] proposed a CNN-based method for detecting cyber attacks in industrial control systems. Reference [58] developed a CNN-based method for detecting web attacks based on HTTP request packets. Ref. [59] utilised a deep learning approach based on CNN-LSTM to detect malware in real time, and the proposed model achieved a high accuracy of 99%.

Reference [60] proposed a DoS detection technique based on LSTM and Bayes and achieved a good performance, according to the authors. References [61,62] used DL techniques for IDS based on the CIC-IDS2018 dataset to improve intrusion detection and CNN and LSTM techniques. In [63], the authors proposed a deep learning technique for detecting cyber attacks. The proposed model used RNN, LSTM and Multilayer Perceptrons (MPs) using a CTT and achieved an accuracy of 93% on LSTM. Reference [64] presented an IDS system based on a deep auto-encoder using the KDD-CUP'99 dataset to evaluate the performance of their model and achieved good results.

3. Proposed Cyber Event Forecasting Model

In this research, a time series-based cyber event forecasting model was proposed. Cyber attacks are sophisticated, and this model aims to help predict cyber attack events. The proactive approaches will help security teams and senior managers implement approaches that protect their network systems. We leveraged a publicly available dataset that contains multiple attack labels collected from a realistic and secure network. The composition of the dataset is Benign 64%, SSH-Bruteforce 18% and FTP-Bruteforce 18%.

Although the dataset was gathered over a five-day period, we concentrated mostly on the 24 h data set that included the three labels. We then performed data sampling using 30 s intervals, resulting in 1,048,575 observations and 79 features. Next, we separated the data into test and training sets. This work is centred on the forecasting of cyber events, with the goal of predicting future attacks that are anticipated to transpire within a given time frame, given a combination of selected events or features. A depiction of the forecasting and other phases of data preparation is given in Figure 1.

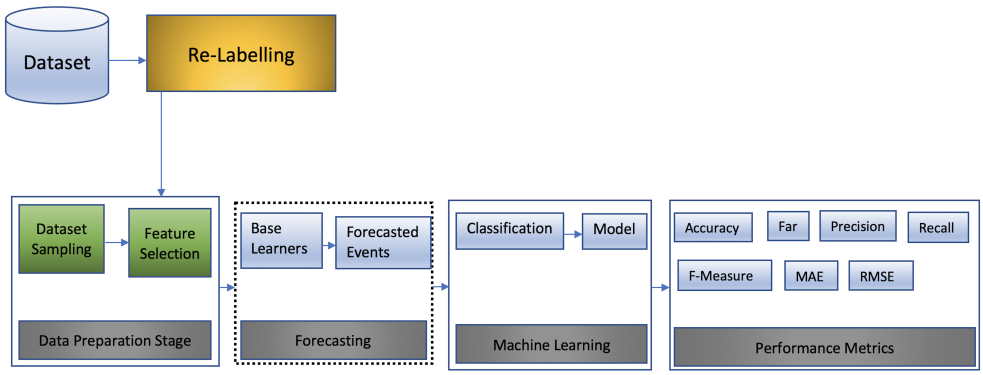

Figure 1. Forecasting stages.

In the next part, we are going to discuss how we prepared the data.

3.1. Data Preparation

The dataset that we utilised was generated from a realistic network and is called CSE-CIC-IDS2018 [17]. As seen in Figure 1, the stages of dataset preparation and experi-

mentation include (i) dataset preparation and feature selection, (ii) classification, (iii) time series forecasting, and (iv) performance evaluation.

The dataset had 1,048,575 observations and 79 features. Data selection was conducted based on a 30 s time interval, which resulted in 1084 observations and 79 features. The time-series data were examined to make sure they were stationary.

3.2. Attack Classification

We applied supervised learning to classify and categorise the observations. The attack classification was performed after the forecasting was completed. We applied popular classification algorithms such as Bayes Net, Naive Bayes, k-NN, Support Vector Machine (SVM) and random forest.

3.3. Feature Selection

Feature selection was performed to determine which of these features were most pertinent to our model. We investigated a number of feature selection techniques, including OneR, Gain Ratio (GR) and Information Gain (IG). Next, using the Information Gain (IG) feature selection method, the top 21 features were chosen. They were chosen based on how well these features ranked and contributed to the model. The portion of the data with the first 884 data points was designated as training and the remaining 200 as the test data.

4. Experiment Setup

The study utilised the CSE-CIC-IDS2018 dataset, and data resampling and time series forecasting were conducted. The primary dataset comprised seven attack categories: (i) Bruteforce, (ii) DoS attack, (iii) Web attack, (iv) Botnet attacks, (v) Infiltration, (vi) DDos and (vii) Heartbleed. Our research focused on a 24 h data subset containing SSH-Bruteforce, FTP-Bruteforce and Benign data. The network under scrutiny encompassed five departments housing 450 computers and 30 servers. Additionally, the attack network comprised 50 machines. The dataset was comprehensive, comprising network traffic and log files from each network host. Data collection occurred at regular 30 s intervals using time series, resulting in 1084 observations. The ratio of benign and cyber events for these observations is 66% and 34%, respectively. Although the cyber events could begin before the first 30 s window and continue after, we aim to perform initial identification of possible cyber attacks by forecasting relevant network traffic features that typically characterise the possibility of attacks. To do so, we look at the past history of the feature values to establish a causal relationship and utilise machine learning models to forecast the future values of the feature. The future values could help inform if an attack is likely to happen and suggest preventative strategies for mitigation.

Subsequently, the dataset was partitioned into training and test segments. Figure 2 illustrates a flow chart delineating the sequential steps and algorithms employed in the experiment. We anticipate that this flow chart will better elucidate the procedural aspects involved in applying machine learning techniques.

4.1. Experiment Overview

The experiment started with a baseline classification that will be used as a reference point for the performance of the forecasted events. We then performed time-series forecasting to evaluate the model. The next parts will cover the results of those experiments.

4.2. Baseline Classification

During the data preparation phase, our dataset was trimmed down to 1084 observations. Subsequently, we allocated 884 observations for baseline training and set aside the remaining 200 for testing purposes. Classification tasks were conducted on the baseline data using BayeNet, Naive Bayes, KNN, SVM and random forest algorithms. This served as a benchmark to assess the accuracy of forecasting cyber events. Table 1 presents the outcomes achieved by these classifiers. The analysis revealed that random forest attained

the highest accuracy at 99.2%, followed closely by KNN and Naive Bayes, which achieved accuracy scores of 98.9% and 98.1%, respectively. To further evaluate the models' performances, we utilised performance metrics. These metrics will be elaborated upon in the evaluation metrics section (Section 5.1).

Table 1. Baseline accuracy scores.

Classifier	Accuracy Score (%)	FAR (%)	Precision (%)	Recall (%)	F-Measure (%)
BayesNet	95.5%	0.5	96.9	95.5	95.8
Naive Bayes	98.1	0.2	98.4	98.1	98.2
k-NN	98.9	0.1	98.9	98.9	98.9
SMO	97.7	0.5	97.9	97.7	97.8
Random Forest	99.2	0.1	99.3	99.2	99.3

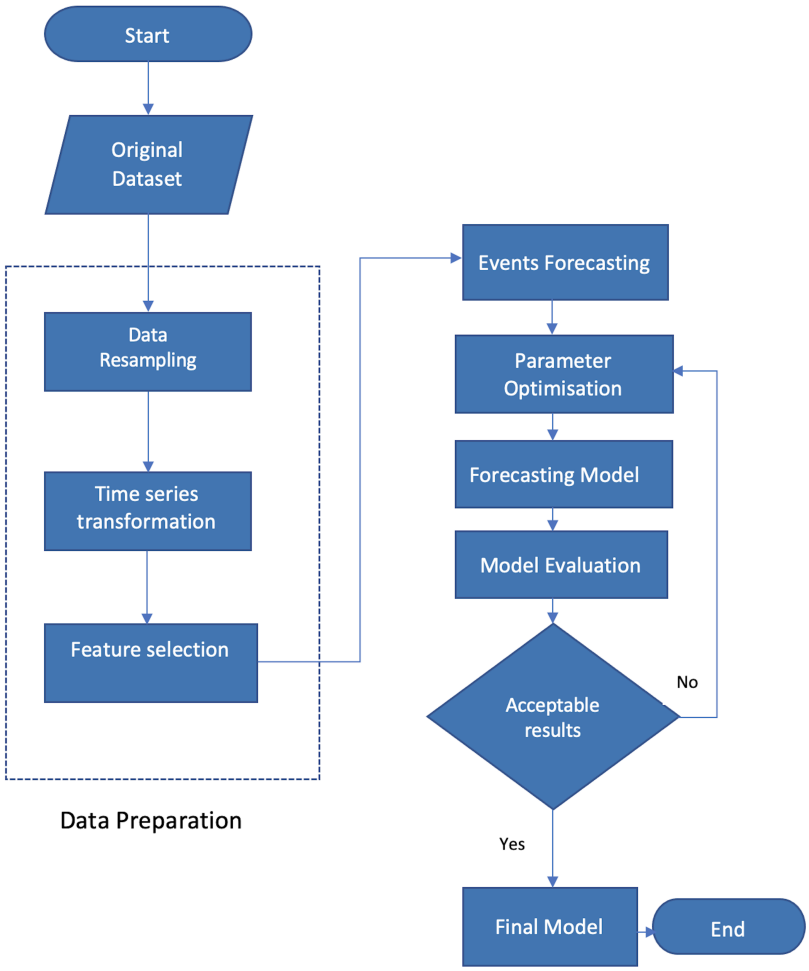

Figure 2. Flow chart—forecasting model.

4.2.1. Classification with Forecasted Values

We used linear regression and SMOreg to perform time series prediction on the forecasted features, totalling 200 observations. We then performed classification on the forecasted events using various classification algorithms, as shown in Tables 2 and 3. We began the classification with the time series events obtained through linear regression. The results show that SVM and random forest were the best-performing classifiers, with an

accuracy score of 90.3% and closely followed by KNN with an accuracy of 89.5%, as shown in Table 2.

Table 2. Classification for forecasted features with linear regression (LNRG).

Base Learner	Classifier	Accuracy Score (%)	FAR (%)	Precision (%)	Recall (%)	F-Measure (%)
LNRG	BayesNet	87.4	1	86.5	87.4	86.3
LNRG	Naive Bayes	87.5	1	86.6	87.5	86.4
LNRG	k-NN	89.5	2	89.9	89.5	88
LNRG	SVM	90.3	2	91.3	90.3	88.9
LNRG	Random Forest	90.3	1	91.3	90.3	88.9

We then repeated the classification using the same algorithms on the time-series events forecasted through SMOreg. The best-performing classification algorithms were SVM and random forest, with an accuracy score of 90.4%, followed by KNN, with an accuracy score of 89.6%. Overall, this accuracy is a slight improvement on the classification accuracy obtained with time-series events derived through linear regression.

Table 3. Classification for forecasted features with Sequential Minimal Optimisation for regression (SMOreg).

Base Learner	Classifier	Accuracy Score (%)	FAR (%)	Precision (%)	Recall (%)	F-Measure (%)
SMOreg	BayesNet	87.5	1	86.6	87.5	86.4
SMOreg	Naive Bayes	87.6	2	86.7	87.6	86.5
SMOreg	k-NN	89.6	2	90	89.6	88.1
SMOreg	SVM	90.4	2	91.3	90.4	89
SMOreg	Random Forest	90.4	1	91.3	90.4	89

Herein, LNRG and SMOreg refer to linear regression and sequential minimal optimisation regression, respectively.

4.2.2. Performance Comparison between Baseline and Forecasted Data

The baseline features had a very high classification accuracy of 99.2% with random forest. In contrast, the forecasted data had the best classification accuracy of 90.4% obtained with SVM and random forest using SMOreg time series data. Overall, the analysis shows a slight drop in performance compared to the classification from the baseline features, but this is to be expected, given these are forecasted events and the limitations posed by the size of the dataset.

Figure 3 shows the classification comparison for Linear Regression (LNRG) and Sequential Minimal Optimisation for regression (SMOreg) forecasted features, the resulted showed SVM and Random Forest performed better with 90/4%.

Figure 4 shows the baseline features' performance accuracy and the forecasted data obtained through linear regression and SMOreg. Although it is a bit lower than the baseline, We believe the 90.4% accuracy score obtained through the time series forecasted features could provide a good prediction accuracy, which will help IT administrators and security professionals be better prepared for cyber incidents and take corrective measures before such attacks materialise.

4.3. Time Series Resampling and Forecasted Events

We conducted time-series forecasting at 30 s intervals, reducing the original dataset from 1,048,575 observations to 1084 observations. Following this, we allocated 884 observations for training purposes, reserving the remaining 200 for testing. The subsequent phase

involved forecasting based on the following 200 events. In this phase of the experiment, we employed linear regression, SMOreg and LSTM models. The subsequent step was to forecast cyber events using these base learners in rotation, followed by assessing the performance of machine learning classification algorithms like Naive Bayes, KNN, SVM and random forest. Additionally, we utilised evaluation metrics such as Mean Absolute Error (MAE) and Root Square Means Error (RSME) to gauge the models' performance.

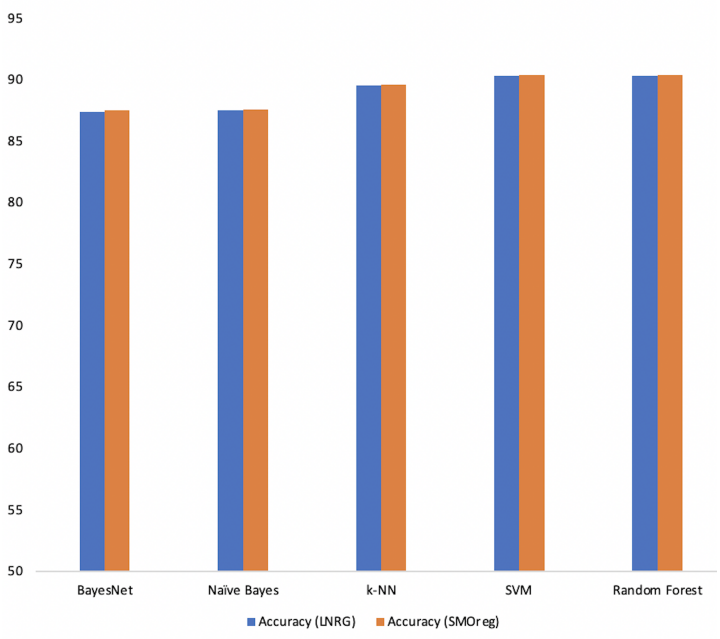

Figure 3. Classification comparison: Linear Regression (LNRG) and Sequential Minimal Optimisation for regression (SMOreg) forecasted features.

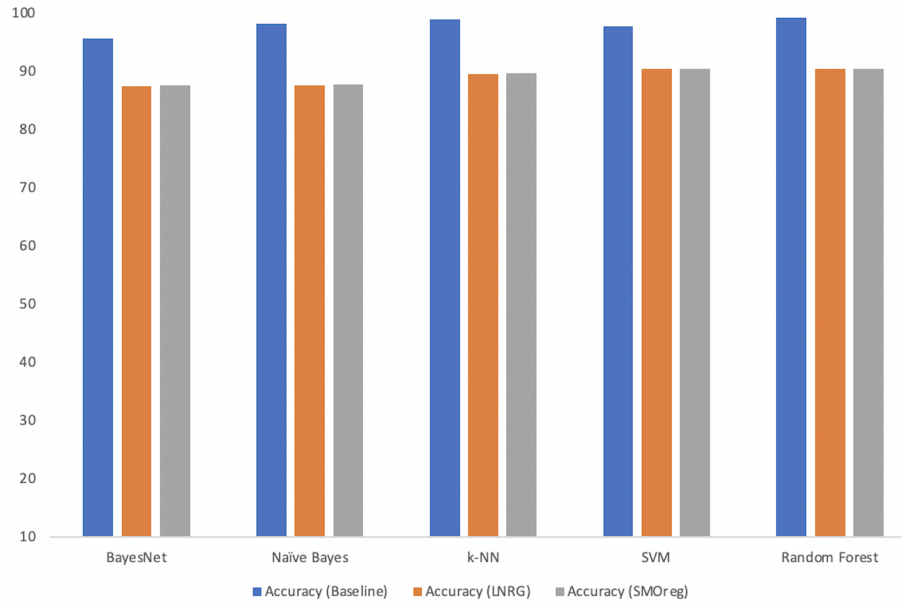

Figure 4. Classification comparison: baseline Linear Regression (LNRG) and Sequential Minimal Optimisation for regression (SMOreg) forecasted features.

4.4. Time Series Forecasting Parameters

The experiment involved utilising the following base learners to forecast cyber events, as detailed in Section 4.3: (i) linear regression, (ii) SMOreg and (iii) LSTM. Subsequently,

the task was to forecast 200 events using each of the aforementioned base learners. The parameters employed in the forecasting experiment are outlined below for each model: (i) attributeSelectionMethods was set to M5 methods, (ii) eliminateColinearAttributes was set to true, (iii) ridge—the default value was selected. In SMOreg, the following parameters were set: (i) the c value = 2.0, (ii) kernel = PolyKernel, (iii) RegOptimizer = RegSMOImproved, (iv) filtetype = normalise training data. In LSTM, the parameters were set to (i) activation function = ActivationReLU, (ii) number of outputs = 3, (iii) gate activation function = ActivationSigmoid. The results of the experiment can be found in Tables 1 and 2.

5. Performance Evaluations

This study employed machine learning classification methods to assess the model's performance. Additionally, we utilised metrics such as Mean Absolute Error (MAE) and Root Square Mean Error (RSME) to evaluate the accuracy of the forecasted data.

5.1. Performance Metrics

During this experiment, we employed various performance metrics to assess our model. These metrics encompass accuracy, precision, recall, F-measure and False Alarm Rate (FAR). Accuracy provides an overall measure of the model's performance, while precision signifies the ratio of correctly classified attacks to instances classified as positive attacks. Recall refers to the correct identification of all relevant instances. Additionally, we used Mean Absolute Error (MAE) and Root Square Mean Error (RSME) to evaluate the accuracy of the forecasted data. The outcomes of the performance metrics are detailed in Tables 1–3. These tables illustrate the classification accuracy of baseline features and forecasted features obtained through linear regression and SMOreg. The baseline features exhibit superior accuracy in evaluation metrics, showcasing a low false alarm rate ranging from 0.1% to 0.5% and a high accuracy of 99.2%. Conversely, the forecasted data display a marginally higher false alarm rate of 1% to 2%. The highest accuracy of 90.4% for the forecasted data was achieved using the SMOreg base learner. The calculation of these metrics is defined in Equations (1)–(5) as utilised in [65].

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$FAR = \frac{FP}{TN + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F - measure(F1) = \frac{2TP}{2TP + FP + FN} \tag{5}$$

True Positive (TP) represents the count of intrusions accurately identified as attacks, while True Negative (TN) signifies the number of regular instances correctly identified as benign packets. False Negative (FN) indicates the quantity of intrusions erroneously labelled as benign packets, whereas False Positive (FP) denotes the quantity of normal instances wrongfully categorized as attacks.

5.2. Mean Absolute Error (MAE)

In this experiment, we utilised the Mean Absolute Error (MAE) metric to assess the performance of the time series forecasted data. Table 4 presents the outcomes derived from the MAE calculations. Within this table, we have highlighted the top five forecasted features acquired through linear regression, SMOreg and LSTM. The findings indicate that the SMOreg forecasted data exhibited superior performance compared to both Linear

Regression and LSTM predictions, showcasing the lowest MAE value. However, it is worth noting that linear regression was relatively close in performance. The equation below illustrates the calculation method for MAE, as referenced in [66].

$$MAE = (\frac{1}{n}) \sum_{i=1}^{n} |y_i - x_i| \quad (6)$$

Herein, y_i = prediction, x_i = actual value and n = total number of data points.

Table 4. Mean absolute error.

Target Feature	Linear Regression	SMOreg	LSTM
Tot Fwd Pkts	0.0105	0.0038	2.6127
Tot Bwd Pkts	0.0034	0.0025	0.4981
Pkt Len Min	0.0002	0.008	0.0054
Fwd Seg Size Min	0.4912	0.0182	0.0494
Subflow Bwd Byts	0.718	0.7251	31.9957

5.3. Root Mean Square Error (RMSE)

The Root Mean Square Error (RMSE) evaluates a model's prediction accuracy by measuring the disparities between actual and predicted values. RMSE computation involves first determining the differences between the numbers, squaring these differences, calculating their mean and then computing the square root of this mean, as depicted in Equation (2) [66]. Within Table 5, we have highlighted the top five forecasted features derived from linear regression, SMOreg and LSTM, assessing their performance based on the RMSE metric. The findings indicate that, in three out of the top five features, the forecasted data from linear regression displayed slightly better performances compared to SMOreg and LSTM predictions, as it exhibited the lowest RMSE value in three out of the top five features.

Table 5. Root mean square error.

Target Feature	Linear Regression	SMOreg	LSTM
Tot Fwd Pkts	0.0129	0.0045	2.8946
Tot Bwd Pkts	0.0039	0.003	0.5842
Pkt Len Min	0.0002	0.0096	0.0063
Fwd Seg Size Min	0.0198	0.8776	0.0497
Subflow Bwd Byts	0.8163	1.2969	35.0395

$$RMSE = \sqrt{(\frac{1}{n}) \sum_{i=1}^{n} (y_i - x_i)^2} \quad (7)$$

5.4. Analysis of the Results

In this experiment, we first determined a baseline containing the original dataset and then performed classifications on them. The baseline classification results showed that random forest had the highest accuracy of 99.2% and a low False Alarm Rate (FAR) of 0.1%, as shown in Table 1. This was closely followed by k-NN, with an accuracy of 98.9%, with SMO, Naive Bayes and BayesNet taking the final spots.

We then performed a classification of the forecasted events. We began the classification of the forecasted features obtained through linear regression techniques. The SVM and random forest forecasted cyber event received the highest classification accuracy of 90.3%,

which is a slight drop compared to the baseline. This was closely followed by k-NN, Naive Bayes and BayesNet with a score of 89.5%, 87.5% and 87.4%, respectively, as shown in Table 2.

We performed classification on the cyber event features forecasted through SMOreg. Again, the SVM and random forest-forecasted cyber event features received the highest classification accuracy of 90.4%, slightly more than those forecasted through linear regression but still lower than the baseline. This was closely followed by k-NN, Naive Bayes and BayesNet with a score of 89.6%, 87.6% and 87.5%, respectively, as shown in Table 3.

Next, we assessed the time series-forecasted events using MAE and RMSE metrics. The findings revealed that SMOreg's forecasted events outperformed those from linear regression and LSTM when evaluating using MAE. Specifically, SMOreg exhibited better predictions in three out of the selected top five features: Tot Fwd Pkts, Tot Bwd Pkts and Fwd Seg Size Min, as illustrated in Table 1. A lower MAE score signifies better performance. Linear regression's forecasted events ranked second, showcasing better performance in two out of the top five features compared to the other two base learners. When evaluated using RMSE, linear regression's forecasted events displayed the highest prediction accuracy in three out of the top five features, contrasting SMOreg's forecasted events, which ranked second in two out of the top five features, as depicted in Table 2. LSTM's performance was inferior to SMOreg and linear regression when assessed using both MAE and RMSE metrics. These metrics serve as crucial indicators for security teams to gauge the model's accuracy and foresee potential cyberattacks, enabling the implementation of preventive measures before the attacks occur.

6. Conclusions and Future Work

In this study, we developed a model for forecasting and predicting cyber events. Utilising a public IDS dataset, we conducted data preparation and resampling. Initial classification on the baseline dataset achieved the highest accuracy score of 99.2% using random forest. Subsequently, cyber event forecasting was carried out by employing linear regression, SMOreg and LSTM to predict the subsequent 200 events for each technique. Machine learning classification algorithms were then applied, with SMOreg-forecasted events yielding the best result at a score of 90.4%. To assess the forecasted events' performance, we utilised the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) metrics. The analysis revealed that SMOreg-forecasted events exhibited superior performance by yielding the lowest MAE in three of the top five selected features. Meanwhile, linear regression's forecasted events demonstrated better performance in three of the top five features when evaluated using the RMSE metric. This study aims to contribute to more accurate cyber attack predictions by anticipating potential attacks based on observed cyber events within the network. It is envisioned that this research will assist security professionals and decision-makers in planning proactive security measures to safeguard systems or critical infrastructure more effectively. The forecasting in this study was confined to specific hourly time frames due to dataset limitations. However, this time window allows security professionals adequate time to respond and proactively address potential threats before they materialise. Future research plans involve expanding this work to encompass the entire dataset spanning five working days, aiming to extend the forecasting window to between 3 and 7 days. It is believed that threats evolve rapidly, and forecasting beyond seven days might render leading indicators obsolete by that time.

Author Contributions: Conceptualization, Y.A.; methodology, Y.A.; validation, Y.A. and T.A.; formal analysis, Y.A.; investigation, Y.A.; and T.A; resources, Y.A., M.A.A. and T.A.; data curation, Y.A.; writing—original draft preparation, Y.A.; writing—review and editing, Y.A., M.A.A. and T.A.; visualization, Y.A.; supervision, Y.A. and T.A.; project administration, Y.A., M.A.A. and T.A.; funding acquisition, Y.A. and M.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was in part funded by the College of Computing, Birmingham City University, UK and Taufiq Asyhari is supported by Monash University-Seed Grant(Indonesia) under grant number IF112304.

Data Availability Statement: The dataset used in this research was from a collaborative project between the Communications Security Establishment (CSE) & the Canadian Institute for Cybersecurity (CIC) and we grateful to them. Available online: https://www.unb.ca/cic/datasets/ids-2018.html (accessed on 1 December 2023).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Constantin, L. Technical Report, CSO Online. 2020. Available online: https://www.csoonline.com/article/3601508/solarwinds-supply-chain-attack-explained-why-organizations-were-not-prepared.html (accessed on 21 Feruary 2023).
2. Dina, A.S.; Siddique, A.; Manivannan, D. A deep learning approach for intrusion detection in Internet of Things using focal loss function. *Internet Things* **2023**, *22*, 100699. [CrossRef]
3. Tang, L.; Li, J.; Du, H.; Li, L.; Wu, J.; Wang, S. Big Data in Forecasting Research: A Literature Review. *Big Data Res.* **2022**, *27*, 100289. [CrossRef]
4. Almahmoud, Z.; Yoo, P.D.; Alhussein, O.; Farhat, I.; Damiani, E. A holistic and proactive approach to forecasting cyber threats. *Sci. Rep.* **2023**, *13*, 8049. [CrossRef] [PubMed]
5. IBM. Cost of a Data Breach 2022. IBM. 2022. Available online: https://www.ibm.com/reports/data-breach (accessed on 11 Feruary 2023).
6. Ghafir, I.; Hammoudeh, M.; Prenosil, V.; Han, L.; Hegarty, R.; Rabie, K.; Aparicio-Navarro, F.J. Detection of advanced persistent threat using machine-learning correlation analysis. *Future Gener. Comput. Syst.* **2018**, *89*, 349–359. [CrossRef]
7. Milajerdi, S.M.; Gjomemo, R.; Eshete, B.; Sekar, R.; Venkatakrishnan, V. Holmes: Real-time apt detection through correlation of suspicious information flows. In Proceedings of the 2019 IEEE Symposium on Security and Privacy (SP), Francisco, CA, USA, 19–23 May 2019 ; IEEE: Piscataway, NJ, USA, 2019; pp. 1137–1152.
8. Ahmed, Y.; Asyhari, A.; Rahman, M.A. A Cyber Kill Chain Approach for Detecting Advanced Persistent Threats. *Comput. Mater. Contin.* **2021**, *67*, 2497–2513. [CrossRef]
9. Laxminarayana, N.; Mishra, N.; Tiwari, P.; Garg, S.; Behera, B.K.; Farouk, A. Quantum-Assisted Activation for Supervised Learning in Healthcare-based Intrusion Detection Systems. *IEEE Trans. Artif. Intell.* **2022**, 1–8. [CrossRef]
10. Gao, Y.; Chen, J.; Miao, H.; Song, B.; Lu, Y.; Pan, W. Self-Learning Spatial Distribution-Based Intrusion Detection for Industrial Cyber-Physical Systems. *IEEE Trans. Comput. Soc. Syst.* **2022**, *9*, 1693–1702. [CrossRef]
11. Abdel Wahab, O. Intrusion Detection in the IoT Under Data and Concept Drifts: Online Deep Learning Approach. *IEEE Internet Things J.* **2022**, *9*, 19706–19716. [CrossRef]
12. Werner, G.; Okutan, A.; Yang, S.; McConky, K. Forecasting Cyberattacks as Time Series with Different Aggregation Granularity. In Proceedings of the 2018 IEEE International Symposium on Technologies for Homeland Security (HST), Woburn, MA, USA, 23–24 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–7.
13. Khandpur, R.P.; Ji, T.; Jan, S.; Wang, G.; Lu, C.T.; Ramakrishnan, N. Crowdsourcing cybersecurity: Cyber attack detection using social media. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, Singapore, 6–10 November 2017; pp. 1049–1057.
14. Hammouchi, H.; Mezzour, G.; Ghogho, M.; El Koutbi, M. Predicting probing rate severity by leveraging twitter sentiments. In Proceedings of the 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Tangier, Morocco, 24–28 June 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 883–888.
15. Goyal, P.; Hossain, K.; Deb, A.; Tavabi, N.; Bartley, N.; Abeliuk, A.; Ferrara, E.; Lerman, K. Discovering signals from web sources to predict cyber attacks. *arXiv* **2018**, arXiv:1806.03342.
16. Tang, M.; Alazab, M.; Luo, Y.; Donlon, M. Disclosure of cyber security vulnerabilities: Time series modelling. *Int. J. Electron. Secur. Digit. Forensics* **2018**, *10*, 255–275. [CrossRef]
17. CSE-CIC. A Realistic Cyber Defense Dataset (CSE-CIC-IDS2018). Technical Report, CSE-CIC. 2018. Available online: https://registry.opendata.aws/cse-cic-ids2018 (accessed on 21 February 2022).
18. Ahmad, Z.; Shahid Khan, A.; Wai Shiang, C.; Abdullah, J.; Ahmad, F. Network intrusion detection system: A systematic study of machine learning and deep learning approaches. *Trans. Emerg. Telecommun. Technol.* **2021**, *32*, e4150. [CrossRef]
19. Zeek. Zeek an Open Source Network Security Monitoring Tool. Available online: https://zeek.org (accessed on 31 March 2023).
20. Snort. Snort Network Intrusion Detection. Available online: https://www.snort.org (accessed on 23 June 2023).
21. Devan, P.; Khare, N. An efficient XGBoost—DNN-based classification model for network intrusion detection system. *Neural Comput. Appl.* **2020**, *32*, 12499–12514. [CrossRef]
22. Gurung, S.; Ghose, M.K.; Subedi, A. Deep learning approach on network intrusion detection system using NSL-KDD dataset. *Int. J. Comput. Netw. Inf. Secur.* **2019**, *11*, 8–14. [CrossRef]
23. Divyasree, T.; Sherly, K. A network intrusion detection system based on ensemble CVM using efficient feature selection approach. *Procedia Comput. Sci.* **2018**, *143*, 442–449. [CrossRef]

24. Splunk. Turn Data into Doing. Available online: https://www.splunk.com (accessed on 31 March 2022).
25. Open Source HIDS. Available online: https://www.ossec.net (accessed on 31 March 2022) .
26. Deshpande, P.; Sharma, S.C.; Peddoju, S.K.; Junaid, S. HIDS: A host based intrusion detection system for cloud computing environment. *Int. J. Syst. Assur. Eng. Manag.* **2018**, *9*, 567–576. [CrossRef]
27. Bulle, B.B.; Santin, A.O.; Viegas, E.K.; dos Santos, R.R. A host-based intrusion detection model based on OS diversity for SCADA. In Proceedings of the IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, Singapore, 18–21 October 2020; pp. 691–696.
28. Chawla, A.; Lee, B.; Fallon, S.; Jacob, P. Host based intrusion detection system with combined CNN/RNN model. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases ; Springer: Dublin, Ireland, 2018; pp. 149–158.
29. Yu, J.; Tao, D.; Lin, Z. A hybrid web log based intrusion detection model. In Proceedings of the 2016 4th International Conference on Cloud Computing and Intelligence Systems (CCIS), Beijing, China, 17–19 August 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 356–360.
30. Otoum, Y.; Nayak, A. As-ids: Anomaly and signature based ids for the internet of things. *J. Netw. Syst. Manag.* **2021**, *29*, 1–26. [CrossRef]
31. Shurman, M.M.; Khrais, R.M.; Yateem, A.A. IoT denial-of-service attack detection and prevention using hybrid IDS. In Proceedings of the 2019 International Arab Conference on Information Technology (ACIT), Al Ain, United Arab Emirates, 3–5 December 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 252–254.
32. Ullah, S.; Khan, M.A.; Ahmad, J.; Jamal, S.S.; e Huma, Z.; Hassan, M.T.; Pitropakis, N.; Buchanan, W.J. HDL-IDS: A hybrid deep learning architecture for intrusion detection in the Internet of Vehicles. *Sensors* **2022**, *22*, 1340. [CrossRef] [PubMed]
33. Singh, V.K.; Govindarasu, M. Cyber Kill Chain-Based Hybrid Intrusion Detection System for Smart Grid. In *Wide Area Power Systems Stability, Protection, and Security*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 571–599.
34. Intrusion Detection Dataset. Available online: https://www.unb.ca/cic/datasets (accessed on 31 March 2022)
35. Traffic Data from Kyoto University's Hotspots. Available online: http://www.takakura.com/Kyoto_data (accessed on 31 March 2022).
36. The UNSW-NB15 Dataset. Available online: https://research.unsw.edu.au/projects/unsw-nb15-dataset (accessed on 31 March 2022)
37. Nagaraj, P.; Krishna, P.S.; Sai, P.S. Forecasting Cyber Attacks Using Machine Learning. *J. Optoelectron. Laser* **2022**, *41*, 550–556.
38. Sadefo Kamdem, J.; Selambi, D. *Cyber-Risk Forecasting Using Machine Learning Models and Generalized Extreme Value Distributions*; Technical Report; HAL: Bengaluru, India, 2022.
39. Deb, A.; Lerman, K.; Ferrara, E. Predicting cyber-events by leveraging hacker sentiment. *Information* **2018**, *9*, 280. [CrossRef]
40. Shu, K.; Sliva, A.; Sampson, J.; Liu, H. Understanding cyber attack behaviors with sentiment information on social media. In Proceedings of the International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation ; Springer: Washington, DC, USA, 2018; pp. 377–388.
41. Alrawi, O.; Ike, M.; Pruett, M.; Kasturi, R.P.; Barua, S.; Hirani, T.; Hill, B.; Saltaformaggio, B. Forecasting Malware Capabilities From Cyber Attack Memory Images. In Proceedings of the USENIX Security Symposium, Virtual, 11–13 August 2021; pp. 3523–3540.
42. Husák, M.; Bartoš, V.; Sokol, P.; Gajdoš, A. Predictive methods in cyber defense: Current experience and research challenges. *Future Gener. Comput. Syst.* **2021**, *115*, 517–530. [CrossRef]
43. Husák, M.; Komárková, J.; Bou-Harb, E.; Čeleda, P. Survey of attack projection, prediction, and forecasting in cyber security. *IEEE Commun. Surv. Tutor.* **2018**, *21*, 640–660. [CrossRef]
44. Sun, N.; Zhang, J.; Rimba, P.; Gao, S.; Zhang, L.Y.; Xiang, Y. Data-driven cybersecurity incident prediction: A survey. *IEEE Commun. Surv. Tutor.* **2018**, *21*, 1744–1772. [CrossRef]
45. Bashar, M.A.; Nayak, R. TAnoGAN: Time series anomaly detection with generative adversarial networks. In Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, Australia, 1–4 December 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1778–1785.
46. Schmidl, S.; Wenig, P.; Papenbrock, T. Anomaly detection in time series: A comprehensive evaluation. *Proc. VLDB Endow.* **2022**, *15*, 1779–1797. [CrossRef]
47. Roumani, Y.; Nwankpa, J.K.; Roumani, Y.F. Time series modeling of vulnerabilities. *Comput. Secur.* **2015**, *51*, 32–40. [CrossRef]
48. Jain, G.; Mallick, B. A study of time series models ARIMA and ETS. Int. J. Mod. Educ. Comput. Sci. **2017**, *4*, 57–63. Available online: http://www.mecs-press.org (accessed on 30 May 2023). [CrossRef]
49. Siami-Namini, S.; Namin, A.S. Forecasting economics and financial time series: ARIMA vs. LSTM. *arXiv* **2018**, arXiv:1803.06386.
50. Werner, G.; Yang, S.; McConky, K. Time series forecasting of cyber attack intensity. In Proceedings of the 12th Annual Conference on Cyber and Information Security Research, Oak Ridge, TN, USA, 4–6 April 2017; pp. 1–3.
51. Liu, H.; Jiang, R.; Zhou, B.; Rong, X.; Li, J.; Li, A. A Survey of Cyber Security Approaches for Prediction. In Proceedings of the 2021 IEEE Sixth International Conference on Data Science in Cyberspace (DSC), ShenZhen, China, 9–11 October 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 439–444.
52. Soundarya, C.; Usha, S. Analyzing and Predicting Cyber Hacking with Time Series Models. *Int. J. Res. Eng. Sci. Manag.* **2020**, *3*, 1–8.

53. Valipour, M. Long-term runoff study using SARIMA and ARIMA models in the United States. *Meteorol. Appl.* **2015**, *22*, 592–598. [CrossRef]
54. Kumari, K.; Yadav, S. Linear regression analysis study. *J. Pract. Cardiovasc. Sci.* **2018**, *4*, 33. [CrossRef]
55. Gabralla, L.A.; Abraham, A. Prediction of oil prices using bagging and random subspace. In Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA 2014, Ostrava, Czech Republic, 23–25 June 2014; pp. 343–354.
56. Coşkun, M.; Yildirim, Ö.; Ayşegül, U.; Demir, Y. An overview of popular deep learning methods. *Eur. J. Tech. (EJT)* **2017**, *7*, 165–176. [CrossRef]
57. Nedeljkovic, D.; Jakovljevic, Z. CNN based method for the development of cyber-attacks detection algorithms in industrial control systems. *Comput. Secur.* **2022**, *114*, 102585. [CrossRef]
58. Zhang, M.; Xu, B.; Bai, S.; Lu, S.; Lin, Z. A deep learning method to detect web attacks using a specially designed CNN. In Proceedings of the Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, 14–18 November 2017; Springer: Cham, Switzerland, 2017; pp. 828–836; Proceedings, Part V 24.
59. Akhtar, M.S.; Feng, T. Detection of Malware by Deep Learning as CNN-LSTM Machine Learning Techniques in Real Time. *Symmetry* **2022**, *14*, 2308. [CrossRef]
60. Li, Y.; Lu, Y. LSTM-BA: DDoS detection approach combining LSTM and Bayes. In Proceedings of the 2019 Seventh International Conference on Advanced Cloud and Big Data (CBD), Suzhou, China, 21–22 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 180–185.
61. Dey, A. Deep IDS: A deep learning approach for Intrusion detection based on IDS 2018. In Proceedings of the 2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 19–20 December 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–5.
62. Farhan, R.I.; Abeer, T.; Nidaa, F. Performance analysis of flow-based attacks detection on CSE-CIC-IDS2018 dataset using deep learning. *Indones. J. Electr. Eng. Comput. Sci.* **2020**, *20*, 16–27. [CrossRef]
63. Ben Fredj, O.; Mihoub, A.; Krichen, M.; Cheikhrouhou, O.; Derhab, A. CyberSecurity attack prediction: a deep learning approach. In Proceedings of the 13th International Conference on Security of Information and Networks, Istanbul, Turkey, 4–6 November 2020; pp. 1–6.
64. Farahnakian, F.; Heikkonen, J. A deep auto-encoder based approach for intrusion detection system. In Proceedings of the 2018 20th International Conference on Advanced Communication Technology (ICACT), Chuncheon-si, Gangwon-do, Republic of Korea, 11–14 February 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 178–183.
65. Aminanto, M.E.; Choi, R.; Tanuwidjaja, H.C.; Yoo, P.D.; Kim, K. Deep abstraction and weighted feature selection for Wi-Fi impersonation detection. *IEEE Trans. Inf. Forensics Secur.* **2017**, *13*, 621–636. [CrossRef]
66. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE). *Geosci. Model Dev. Discuss.* **2014**, *7*, 1525–1534.