Global Network Cyberattack Classification Using Naive Bayes Method Time Range 2020 – 2023

Acep Sandi Mutia, Irawan Irawan, Christina Juliane

STMIK Likmi, Bandung, INDONESIA

E-mail: acepsandimutia707@gmail.com, irawan.skom@gmail.com, christina.juliane@likmi.ac.id

Submitted: December 18, 2023 | Revised: December 19, 2023 | Accepted: May 20, 2024 |
Published: May 22, 2024

ABSTRACT

This study focuses on developing a classification model for cyberattacks on global networks during the time span of 2020 to 2023 using the Naive Bayes method. The main objective of the study is to analyze and classify the frequent severity of cyber, which helps in improving network security and reducing vulnerabilities. The Naive Bayes method was chosen for its efficiency in handling large datasets and its ability to make predictions based on probabilities. Collecting cyberattack data from a variety of reliable and up-to-date sources, the study covers attacks such as ransomware, phishing, DDoS, and other malware. The classification process includes data pre-processing, feature extraction, and finally the application of Naive Bayes algorithms to identify patterns in such attacks. The classification results are then evaluated using the Apply Model and Performance validation methods to assess the effectiveness of the model. The results of this study show that Naive Bayes is able to accurately classify cyberattacks, providing a useful tool for cybersecurity professionals to understand attack trends and respond proactively. The study also suggests areas for further research, including the integration of the Naive Bayes model with other artificial intelligence systems for improved cyberattack detection. The study provides new insights into the application of the Naive Bayes method in cybersecurity and paves the way for improved data-driven cyber defense strategies.

Keywords: data mining; classification of cyberattacks; naive bayes; network security; security data analysis.

INTRODUCTION

In today's digital age, cybersecurity has become a top priority for both individuals, organizations, because the security of every digital device that is part of a global network with millions of computer nodes is critical (Veeramanickam et al., 2022). The rapid growth in information and communication technology has opened up opportunities for global economic activity. However, on the other hand, it also poses opportunities for different types of cyberattacks. From ransomware to DDoS (Distributed Denial of Service) attacks. DDoS attacks flood a network with traffic from many infected computers to disrupt or disable access to a target's device or service (Shukla et al., 2023).

The years 2020 to 2023 recorded a significant increase in the number and complexity of cyberattacks. The COVID-19 pandemic, for example, became a catalyst for cyberattacks as it transitioned massively to remote work, widening the attack surface for cybercriminals. The sophistication of AI and machine learning technology has also been leveraged by cyber attack actors to create threats that are more adaptive and difficult to detect (Amin et al., 2020; Bécue et al., 2021; Kuhn et al., 2021).

Classifying cyberattacks is critical because it helps in identifying, analyzing, and handling threats more effectively. Classification allows cybersecurity experts to recognize patterns and distinctive features of different types of attacks, which can speed up the process of detection and response to security incidents. With a better understanding of the nature and characteristics of attacks, security strategies can be tailored to be more resilient in the face of specific threats (Lin et al., 2022; Syrmakesis et al., 2022).

At the global level, effective classification of cyberattacks enables better collaboration and coordination among countries and international organizations in the fight against cybercrime. This is important given that cyberattacks are often transnational in nature, crossing national borders and
legal jurisdictions. Therefore, international cooperation, supported by a deep understanding of cyberattacks, is key to responding effectively to cyber threats and minimizing their impact on global security (Sarker, 2023; Sarker et al., 2020).

Through efficient classification, we can also understand the evolution of cyber threats and anticipate future trends. This helps in the development of cybersecurity solutions that are proactive, rather than just reactive, ensuring that security measures remain relevant and effective in the face of evolving threats (Abu Al-Haija & Al-Fayoumi, 2023).

Overall, a thorough understanding of cyberattacks and accurate classification of these threats are critical components in a global cybersecurity strategy, ensuring data security, critical infrastructure, and maintaining the integrity of information systems around the world.

The Naive Bayes method is a statistical classification technique based on Bayes' Theorem. This is one of the most simple and effective machine learning models, especially in the case of large-dimensional data. Naive Bayes operate under the assumption that features in the dataset are independent of each other, an approach known as "naivety". Although this assumption is sometimes unrealistic in practice, Naive Bayes proved efficient in a wide range of practical applications, especially in natural language processing and text classification (Kim & Lee, 2022).

In the context of cyberattack classification, Naive Bayes was chosen for several reasons. First, its ability to handle large amounts of data closely matches the vast and diverse characteristics of cyberattack data. Secondly, this method is efficient in terms of computational time, which is very important in the detection of cyberattacks, where response speed is key. Third, Naive Bayes can operate well even using datasets that have noise and incomplete features, which often occur in cyberattack data (Shi et al., 2021).

In the classification of cyberattacks, Naive Bayes are used to identify and categorize different types of attacks based on features present in the data, such as network traffic patterns, payload types, and other abnormal behaviors. The model is trained with a dataset that includes known instances of attacks, allowing the algorithm to 'learn' the different patterns and characteristics of each type of attack.

The main advantages of Naive Bayes are ease of implementation and efficiency in computing. In addition, these models tend to work well even in imperfect data conditions. However, there are also some challenges. Assumptions of feature independence are often unrealistic in cyberattack data, where features can be interrelated. In addition, Naive Bayes can be less effective if the distribution of data within classes is not uniform or if there are dominant features that can give rise to bias (Chen et al., 2021; Chu et al., 2020; Redivo et al., 2023).

Overall, despite its limitations, Naive Bayes remains a popular choice in the classification of cyberattacks due to its practicality and efficiency. Especially in situations where speed and the ability to manage large volumes of data are a priority (Blanquero et al., 2021).

Information systems that support network planning activities are technology platforms used to collect, process and analyze data to assist in the network planning and management process. The system integrates various types of data, including geographic, demographic and infrastructure data, to provide a comprehensive picture of field needs and conditions (Sari OL et.al, 2024; Nur A et.al, 2024).

In the data processing process, this system carries out needs analysis and forecasting to predict future network demand. Modeling and simulation are used to test various network design scenarios and select the best option based on performance, cost, and scalability. Furthermore, this information system supports network topology design by determining the optimal structure and allocating resources efficiently. Project management processes, including scheduling, monitoring, and budget management, are also facilitated by this system to ensure network implementation runs smoothly and according to plan. Additionally, the system provides reporting and documentation tools to monitor network performance and manage technical documentation. Collaboration and communication between teams is made easy through collaboration portals and real-time
communication tools, while data security is maintained through encryption and compliance with privacy regulations (Pradnyana IM et al., 2023; Prastowo FI et al., 2023).

**RESEARCH METHOD**

This study applies algorithm modeling techniques using the Knowledge Discovery in Database (KDD) approach. The purpose of this modeling technique is to extract previously unknown information and understanding from the database. The explanation can be seen in figure 1 (Fahd et al., 2022).

1. **Collecting Data**
   Collecting data is the stage of data that has been collected. The data used is sourced from kaggle.com, which is open-source (Schoenenwald et al., 2021).

2. **Cleansing**
   The collected data will be cleaned and some will be deleted. This process includes cleaning bad data, data that has empty attributes, abnormal data, as well as unused attributes when modeling. In this dataset we cleanse ip addresses that do not give a ping response (Hosseinzadeh et al., 2021).

3. **Transformation**
   Transformations are used to produce optimal performance in data modeling, some data that is not directly related is reduced to improve accuracy. The data transformation process is carried out on the dataset so that the data can be used in this research when modeling. This process can also have an effect on the modeling results displayed at the evaluation stage (Damayunita et al., 2022).

4. **Modeling**
   Modeling After completing the data cleansing and transformation phase, the modeling stage is run. At this stage, the results of classification and predictions are established. In this study, the Bayesian Naive modeling algorithm was used (Damayunita et al., 2022).

5. **Evaluation**
   The results of modeling experiments are displayed in the form of a fusion matrix or error matrix. This fusion matrix represents actual information about the modeling performed and also provides information in the form of accuracy results (Damayunita et al., 2022).

![Figure 1 Method](image-url)

**RESULTS AND DISCUSSION**

**Collecting Data**

At this stage, data collected from kaggle.com about the cyber attack global network between 2022 to 2023 amounted to 40,000 data records and 25 attributes. The explanation of each attribute can be seen in table number 1, while the sample dataset can be seen in figure 2 with the attribute name placed in column 2 so that the example record can be read properly.
Table 1. Attribute Description

<table>
<thead>
<tr>
<th>No</th>
<th>Attribute Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>timestamps</td>
<td>Logging date</td>
</tr>
<tr>
<td>2</td>
<td>source_ip_address</td>
<td>The origin of the IP Address that carried out the attack</td>
</tr>
<tr>
<td>3</td>
<td>destination_ip_address</td>
<td>IP Address korban serangan</td>
</tr>
<tr>
<td>4</td>
<td>source_port</td>
<td>Asal port</td>
</tr>
<tr>
<td>5</td>
<td>destination_port</td>
<td>Purpose of the attacked port</td>
</tr>
<tr>
<td>6</td>
<td>protocol</td>
<td>Protocol type</td>
</tr>
<tr>
<td>7</td>
<td>packet_length</td>
<td>Large data packets sent</td>
</tr>
<tr>
<td>8</td>
<td>packet_type</td>
<td>Package type</td>
</tr>
<tr>
<td>9</td>
<td>traffic_type</td>
<td>Traffic by type</td>
</tr>
<tr>
<td>10</td>
<td>payload_data</td>
<td>Fill in the data sent</td>
</tr>
<tr>
<td>11</td>
<td>malware_indicators</td>
<td>Malware indicator</td>
</tr>
<tr>
<td>12</td>
<td>anomaly_scores</td>
<td>Nilai anomaly way</td>
</tr>
<tr>
<td>13</td>
<td>alerts_warnings</td>
<td>Automatic Alerts are called</td>
</tr>
<tr>
<td>14</td>
<td>attack_type</td>
<td>Types of attacks</td>
</tr>
<tr>
<td>15</td>
<td>attack_signature</td>
<td>Attack identifier</td>
</tr>
<tr>
<td>16</td>
<td>action_taken</td>
<td>Interceptor steps of attacks performed</td>
</tr>
<tr>
<td>17</td>
<td>severity_level</td>
<td>Attack damage level</td>
</tr>
<tr>
<td>18</td>
<td>user_information</td>
<td>User information</td>
</tr>
<tr>
<td>19</td>
<td>device_information</td>
<td>Application information used</td>
</tr>
<tr>
<td>20</td>
<td>network_segment</td>
<td>Network segments</td>
</tr>
<tr>
<td>21</td>
<td>Geolocation_data</td>
<td>Where the attack came from</td>
</tr>
<tr>
<td>22</td>
<td>proxy_information</td>
<td>Proxy information</td>
</tr>
<tr>
<td>23</td>
<td>firewall_logs</td>
<td>Firewall notes</td>
</tr>
<tr>
<td>24</td>
<td>idsips_alerts</td>
<td>IDS record</td>
</tr>
<tr>
<td>25</td>
<td>log_source</td>
<td>Logging origin</td>
</tr>
</tbody>
</table>

Figure 2. Datasets displayed horizontally
Cleaning Data

Once data collection is complete, the next step is data cleansing. At this stage, the data that is considered invalid is eliminated, bringing the total number of data to 8751 data to be retrieved. Of the 8751 data, 80% will be used as a training dataset and 20% will be used as a test dataset. The following script is used to perform the data cleaning process, created using PHP and run in the Linux console prompt.

```php
<?php
    $ip_address = '192.168.1.1'; // Host
    $sql = "SELECT * FROM cyberattack_cleansing ORDER BY ip_address";
    $result = pg_query($sql);
    if ($result) {
        $ip = pg_fetch_assoc($result);
        if ($ip['ip_address'] !== $ip_address) {
            echo "Error: IP address not found.
        } else {
            echo "Record found.
        }
    }
    
    $sql = "UPDATE cyberattack SET online = '1' where ip_address = ".$ip['ip_address'].""
    $result = pg_query($sql);
    if ($result) {
        echo "Record updated successfully.
    } else {
        echo "Error updating record.
    }

    $sql = "DELETE FROM cyberattack WHERE ip_address = ".$ip['ip_address'].""
    $result = pg_query($sql);
    if ($result) {
        echo "Record deleted successfully.
    } else {
        echo "Error deleting record.
    }

    $sql = "SELECT * FROM cyberattack WHERE ip_address = ".$ip['ip_address'].""
    $result = pg_query($sql);
    if ($result) {
        $row = pg_fetch_assoc($result);
        echo "Record found:
        ".
    } else {
        echo "Record not found.
    }

    $sql = "UPDATE cyberattack SET online = '0' where ip_address = ".$ip['ip_address'].""
    $result = pg_query($sql);
    if ($result) {
        echo "Record updated successfully.
    } else {
        echo "Error updating record.
    }
```

Figure 3. Cleansing

From the cleansing program shown in figure 3 after running, all records that have IP addresses that cannot be pinged will be deleted from the dataset, so that the dataset we use is only records whose IP addresses are active at the time this study is made. An example of the dataset we used can be seen in Figure 4.
Acep Sandi Mutia, Irawan Irawan, Christina Juliane
Global Network Cyberattack Classification Using Naive Bayes Method Time Range 2020 – 2023

Figure 4. Results Dataset Cleansing

Transformation Data

The data transformation process establishes a performance evaluation of the algorithms used to convert data, with the features used as encoders being destructive_power, severity_t, attack_t, action_t, malware_indicators_t, and anomaly_scores.

Table 2. Data Transformation

<table>
<thead>
<tr>
<th>Feature</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>destruct_power_character</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>varying</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>severity_t</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>integer</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>attack_t</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>integer</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>action_t</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>integer</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>malware_indicators_t</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>integer</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>anomaly_scores</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>numeric(5,2)</td>
<td>7.26</td>
<td>65.92</td>
<td>12.66</td>
</tr>
</tbody>
</table>

Modeling

This stage aims to find the classification and prediction results from Syber. This study used Naive Bayes algorithm modeling. The formula used to describe Bayes' Naive theorem is listed in formula (1).

\[
P(H|X) = \frac{p(X|H)p(H)}{p(X)}
\] (1)

Information:
X = Data whose class is still not identified.
H = Hypothesis that data X belongs to a certain class.
p(H|X) = Possible hypothesis H taking into account condition X.
p(X|H) = The probability of X given the conditions present in hypothesis H.
p(H) = Possible of hypothesis H.
p(X) = Possible occurrence of X.

Figure 5. Dataset Attack
From the results of data transformation, it is continued by entering the dataset in Figure 5 into the datamining design process in the rapidminer application as shown in figure 6.

Figure 6. Datamining Design

Evaluation
The next step in determining the accuracy value of modeling carried out on test data is to test the test data that has been separated before.

Figure 7. Naïve Baiyes Classification Results
The results of execution in datamining design can be seen in figure 8 which shows the accuracy of this study of 84.53% with detailed results as follows:

- Predicted High and Turned High as much as 450
- High prediction and it turns out Medium as much as 0
- High prediction and it turns out Low as much as 0
- Medium prediction and it turns out High as much as 41
- Medium prediction and it turns out Medium is 815
- Medium prediction and it turns out Low as much as 156
- Low predicted results and it turned out to be High as much as 90
- Low predicted results and it turned out to be Medium as much as 119
- Low's prediction results and it turned out that Low was 954

Table 3: Value Confidence

<table>
<thead>
<tr>
<th>No</th>
<th>attack_n</th>
<th>action_n</th>
<th>severity_n</th>
<th>anomaly_scores</th>
<th>destructive_power</th>
<th>Confidence (High)</th>
<th>Confidence (Medium)</th>
<th>Confidence (Low)</th>
<th>Prediction (destructive_power)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>59.5</td>
<td>High</td>
<td>0.7</td>
<td>0.2</td>
<td>0.0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>61.2</td>
<td>High</td>
<td>0.7</td>
<td>0.2</td>
<td>0.0</td>
<td>High</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>92.2</td>
<td>Medium</td>
<td>0.9</td>
<td>0.0</td>
<td>0.1</td>
<td>Medium</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>56.5</td>
<td>Medium</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>Medium</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>89.6</td>
<td>Low</td>
<td>0.9</td>
<td>0.0</td>
<td>0.1</td>
<td>Medium</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>59.3</td>
<td>Low</td>
<td>0.6</td>
<td>0.0</td>
<td>0.4</td>
<td>Medium</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>9.3</td>
<td>Medium</td>
<td>0.3</td>
<td>0.0</td>
<td>0.7</td>
<td>Low</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>63.4</td>
<td>Medium</td>
<td>1.0</td>
<td>0.0</td>
<td>0.0</td>
<td>Medium</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>49.0</td>
<td>Low</td>
<td>0.2</td>
<td>0.0</td>
<td>0.8</td>
<td>Low</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 describes statistical data from the application of the Naïve Bayes algorithm for the calculation of prediction results on destructive_power labels from attack_n, action_n, severity_n, and anomaly_scores data, for example High's accurate prediction results are highly dependent on the high value of attack_n, action_n, severity_n and anomaly_score data, which shows a considerable damage impact on the system if this data is of high value. And vice versa the impact of damage is low if the data is attack_n, action_n, severity_n and anomaly_score of low value.

CONCLUSION

From the results of this study, the author can draw several conclusions as follows: 1) the use of the Naïve Bayes classification model has proven to be effective in classifying cyberattack datasets with an accuracy rate of 84.53%. 2) the level of accuracy in testing with high prediction results reached 77.45%, Medium prediction results reached 87%, Low prediction results reached 85.95%. The results confirm the accuracy of Naïve Bayes in classifying cyberattacks and recommend their combination with other methods for a more effective cyber defense strategy.

REFERENCES


Chu, S.-C., Dao, T.-K., Pan, J.-S., & Nguyen, T.-T. (2020). Identifying correctness data scheme for...


