PREDICTIVE CYBER INSURANCE POLICY ANALYSIS: ENHANCING CYBER SECURITY MANAGEMENT

Ritesh Kumar[1], Rohit Nandi[2], Rushitha Bommisetty[2], Rakshitha Yadav Marla[2]
[1]Assistant.Professor, [2]UG Scholar, [1,2]Department of Computer Science & Engineering.
[1,2]Kommuri Pratap Reddy Institute of Technology, Ghatkesar, Hyderabad, Telangana.

ABSTRACT:
In today's world, the reality of modern cyberattacks and their severe impacts has become apparent, highlighting the fact that relying solely on risk mitigation measures is not enough for organizational cybersecurity management. As a result, cyber insurance has emerged as a necessary complement to existing safeguards. Some notable cybersecurity attacks with critical severity include WannaCry and NotPetya in 2017, which wreaked havoc on thousands of companies across various regions and industries. Additionally, there was a ransomware attack that affected major governmental organizations in the USA, such as the Departments of Defense, Homeland Security, State, Treasury, Energy, Commerce, and others. These incidents underscore the urgency of bolstering cybersecurity defenses. Today, the digital landscape is filled with advanced cyber threats of high severity, including crypto jacking, malware, supply-chain attacks, ransomware, business email compromise, and more. In this context, cyber insurance has gained increasing importance as organizations face the ever-growing menace of cyberattacks and data breaches. To address this critical issue, accurate prediction of cyber insurance policy patterns can play a vital role. By predicting these patterns, insurance companies can better assess risk, set appropriate premiums, and design effective coverage strategies. To achieve this, a novel methodology is proposed in this work, combining two powerful techniques: TF-IDF (Term Frequency-Inverse Document Frequency) feature extraction and a multinomial naive Bayes classifier. The TF-IDF algorithm is utilized to represent policy documents as numerical feature vectors, which capture the significance of terms within the documents. Subsequently, the multinomial naive Bayes classifier is employed to classify the policy patterns based on the extracted features. This approach presents a promising way to enhance cybersecurity management through predictive cyber insurance policy analysis. By leveraging advanced techniques and algorithms, organizations can better prepare themselves for potential cyber threats, making informed decisions to safeguard their interests and assets.

Keywords: Cyberattacks, Risk mitigation, Malware, Ransomware, TF-IDF, Classification.

1. INTRODUCTION

The history of cyber insurance policy analysis traces back to the early days of the internet, when businesses started grappling with the implications of cybersecurity threats. In the late 20th century, as the internet became increasingly integrated into business operations, companies began to realize the potential financial losses associated with cyberattacks and data breaches. The first cyber insurance policies emerged in the late 1990s, offering coverage for losses related to hacking, viruses, and other cyber incidents. Initially, cyber insurance policies were relatively simplistic, covering basic risks such as data breaches and denial-of-service attacks. However, as cyber threats evolved in complexity and severity, the insurance industry adapted accordingly. Major

cybersecurity incidents, such as the Code Red and Nimda worms in the early 2000s, served as wake-up calls for both businesses and insurers, highlighting the need for more comprehensive coverage. The landscape shifted dramatically in the 2010s with the proliferation of ransomware attacks and nation-state cyber warfare. High-profile incidents like the WannaCry and NotPetya outbreaks in 2017 demonstrated the devastating potential of cyber threats, prompting organizations to reevaluate their cybersecurity strategies. Consequently, demand for cyber insurance surged, leading insurers to develop more sophisticated policies tailored to the evolving threat landscape.

Today, cyber insurance is an integral component of risk management for organizations across industries. Insurers offer a wide range of coverage options, including financial protection against data breaches, business interruption losses, and liability claims arising from cyber incidents. As cyber threats continue to evolve, the history of cyber insurance policy analysis serves as a testament to the importance of staying ahead of emerging risks and adapting insurance strategies accordingly.

2. LITERATURE SURVEY

Kure et al. [1] proposed a novel integrated cyber security risk management (i-CSRM) framework that responds to that challenge by supporting systematic identification of critical assets through the use of a decision support mechanism built on fuzzy set theory, by predicting risk types through machine learning techniques, and by assessing the effectiveness of existing controls. The framework is composed of a language, a process, and it is supported by an automated tool. The paper also reported on the evaluation of our work to a real case study of a critical infrastructure. The results revealed that using the fuzzy set theory in assessing assets' criticality, our work supports stakeholders towards an effective risk management by assessing each asset's criticality. Furthermore, the results have demonstrated the machine learning classifiers' exemplary performance to predict different risk types including denial of service, cyber espionage and crimeware.

Albasheer et al. [2] reviewed the state-of-the-art cyber-attack prediction based on NIDS Intrusion Alert, its models, and limitations. The taxonomy of intrusion alert correlation (AC) is introduced, which includes similarity-based, statistical-based, knowledge-based, and hybrid-based approaches. Moreover, the classification of alert correlation components was also introduced. Alert Correlation Datasets and future research directions are highlighted. The AC received raw alerts to identify the association between different alerts, linking each alert to its related contextual information and predicting a forthcoming alert/attack. It provides a timely, concise, and high-level view of the network security situation. This review can serve as a benchmark for researchers and industries for Network Intrusion Detection Systems' future progress and development.

Tsohou et al. [3] examined the relevant literature on cybersecurity insurance, research and practice, in order to draft the current landscape and present the trends. This has led to an increase of cyberattacks, as a direct consequence of the increase of the attack surface but subsequently also led to an increased necessity for the protection of information systems. Toward the protection of information systems, cyber insurance is considered as a strategy for risk management, where necessary. Cyber insurance is emerging as an important tool to protect organizations against cyberattack-related losses.

Zhao et al. [4] proposed CTP-DHGL, a novel Cyber Threat Prediction model based on Dynamic Heterogeneous Graph Learning, to predict the potential cyber threats by investigating public security-related data (e.g., CVE

details, ExploitDB). Particularly, we first characterize the interactive relationships among different types of cyber threat objects with a heterogeneous graph. This work then formalized cyber threat prediction as a dynamic link prediction task on the heterogeneous graph and propose an end-to-end dynamic heterogeneous graph embedding method to learn the dynamic evolutionary patterns of the graph. As a result, CTP-DHGL can infer potential link relationships based on the evolving graph embedding sequences learned from previous snapshots to infer stealthy cyber threats. The experimental results on real-world datasets verify that CTP-DHGL outperforms the baseline models in learning the evolutionary patterns of cyber threats and predicting potential cyber risks.

Husák et al. [5] studied the both methods based on discrete models, such as attack graphs, Bayesian networks, and Markov models, and continuous models, such as time series and grey models, are surveyed, compared, and contrasted. This work further discussed machine learning and data mining approaches, that have gained a lot of attention recently and appears promising for such a constantly changing environment, which is cyber security. The survey also focused on the practical usability of the methods and problems related to their evaluation.

Singh et al. [6] first look at the soft spots and threats faced by the insurance companies, and the impacts of these threats. This work finds that both management and technology measures are necessary to tackle the threat. This work then come up with a five-pronged recommendation framework on how insurance companies can strengthen their security infrastructure.

Hwang et al. [7] studied the latent Dirichlet allocation is applied to extract text-document-based technical topics for the symmetrical thesis and patent information to identify security convergence fields and technologies for cyber safety. In addition, it elucidates cyber security convergence fields and technology trends by applying a dynamic topic model and long short-term memory, which are useful for analyzing technological changes and predicting trends. Based on these results, cyber security administrators, system operators, and developers can effectively identify and respond to trends in related technologies to reduce threats, and companies and experts developing cyber security solutions can present a new security approach.

Sarker et al. [8] presented an Intrusion Detection Tree ("IntruDTree") machine-learning-based security model that first considers the ranking of security features according to their importance and then build a tree-based generalized intrusion detection model based on the selected important features. This model is not only effective in terms of prediction accuracy for unseen test cases but also minimizes the computational complexity of the model by reducing the feature dimensions. Finally, the effectiveness of our IntruDTree model was examined by conducting experiments on cybersecurity datasets and computing the precision, recall, fscore, accuracy, and ROC values to evaluate. This work also compared the outcome results of IntruDTree model with several traditional popular machine learning methods such as the naive Bayes classifier, logistic regression, support vector machines, and k-nearest neighbor, to analyze the effectiveness of the resulting security model.

Lu et al. [9] established a kind of network safety situation forecast model based on Grey Wolf Optimization (GWO) algorithm to optimize support vector machine (SVM) parameters and solves the problem of support vector machine (SVM) parameter optimization. It overcomes the problems of neural network training and local optimization, which makes it more generalized, also effectively improve the prediction effect of SVM. The simulation experiments indicated that this model has improved the accuracy of prediction and shows the general tendency of the network security situation.

3. PROPOSED METHODOLOGY

A cyber insurance policy, also known as cyber risk insurance or cyber liability insurance, is a type of insurance coverage designed to protect individuals, businesses, and organizations from financial losses and liabilities associated with cyber-related incidents and data breaches. With the increasing frequency and sophistication of cyber-attacks, cyber insurance has become an important risk management tool for many entities.

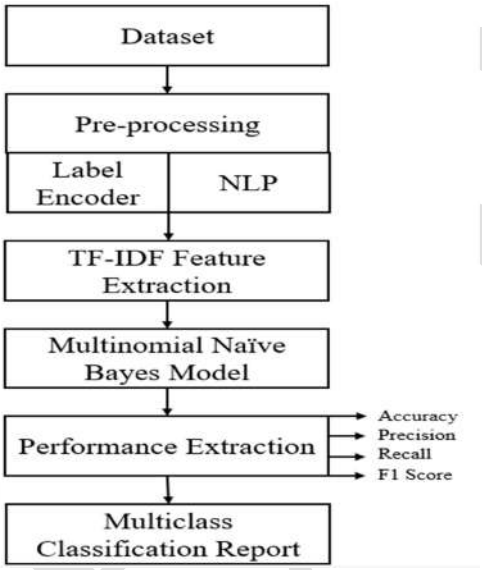

Fig. 1: Block diagram of proposed system.

To predict cyber insurance policy patterns using TF-IDF (Term Frequency-Inverse Document Frequency) and a Multinomial Naïve Bayes classifier, this work following steps:

- Data Preprocessing: Gather a dataset of cyber insurance policy documents. Preprocess the data by removing any irrelevant information, such as headers, footers, or special characters. Tokenize the text into individual words or n-grams (sequences of adjacent words). Perform other preprocessing steps like stemming, lemmatization, or removing stop words based on your specific requirements.
- TF-IDF Vectorization: Convert the pre-processed text documents into numerical features using the TF-IDF technique. TF-IDF assigns weights to words based on their frequency in a document (TF) and inverse frequency across all documents (IDF). This technique helps capture the importance of words within individual documents relative to the entire corpus.
- Splitting the Dataset: Split your dataset into training and testing sets. The training set will be used to train the Multinomial Naïve Bayes classifier, while the testing set will be used to evaluate its performance.
- Training the Classifier: Train a Multinomial Naïve Bayes classifier using the TF-IDF vectors from the training set. The Multinomial Naïve Bayes classifier is suitable for text classification tasks as it assumes that the features (TF-IDF values) are generated from a multinomial distribution.

— Model Evaluation: Evaluate the trained classifier using the testing set. Calculate metrics such as accuracy, precision, recall, and F1-score to assess the performance of the classifier.

TF-IDF Feature Extraction

TF-IDF, short for Term Frequency-Inverse Document Frequency, is a commonly used technique in NLP to determine the significance of words in a document or corpus. To give some background context, a survey conducted in 2015 showed that 83% of text-based recommender systems in digital libraries use TF-IDF for extracting textual features. That's how popular the technique is. Essentially, it measures the importance of a word by comparing its frequency within a specific document with the frequency to its frequency in the entire corpus. The underlying assumption is that a word that occurs more frequently within a document but rarely in the corpus is particularly important in that document.

Mathematical formula for calculating TF-IDF

TF (Term Frequency) is determined by calculating the frequency of a word in a document and dividing it by the total number of words in the document.

— TF = (Number of times the word appears in the document) / (Total number of words in the document)

— IDF (Inverse Document Frequency), on the other hand, measures the importance of a word within the corpus as a whole. It is calculated as:

— IDF = log((Total number of documents in the corpus) / (Number of documents containing the word))

Multinominal Naïve Bayes

Multinomial Naive Bayes algorithm is a probabilistic learning method that is mostly used in Natural Language Processing (NLP). The algorithm is based on the Bayes theorem and predicts the tag of a text such as a piece of email or newspaper article. It calculates the probability of each tag for a given sample and then gives the tag with the highest probability as output.

Naive Bayes classifier is a collection of many algorithms where all the algorithms share one common principle, and that is each feature being classified is not related to any other feature. The presence or absence of a feature does not affect the presence or absence of the other feature.

Naive Bayes is a powerful algorithm that is used for text data analysis and with problems with multiple classes. To understand Naive Bayes theorem's working, it is important to understand the Bayes theorem concept first as it is based on the latter.

Bayes theorem, formulated by Thomas Bayes, calculates the probability of an event occurring based on the prior knowledge of conditions related to an event. It is based on the following formula:

$P(A|B) = P(A) * P(B|A)/P(B)$

Where we are calculating the probability of class A when predictor B is already provided.

P(B) = prior probability of B

P(A) = prior probability of class A

P(B|A) = occurrence of predictor B given class A probability

This formula helps in calculating the probability of the tags in the text.

4. RESULTS

Figure 2 presents a visual representation of the original dataset used in the predictive cyber insurance policy analysis. It shows a tabular layout with rows and columns, where each row corresponds to a specific cyber incident and each column represents a different attribute or feature related to that incident.

	ID	year	Actor	victim.victim_id	org_size	victim.industry.name	pattern	summary
0	137	2015	External	American Tooling Center	11 - 100	Manufacturing	Social Engineering	a federal appeals court has sided with a comme...
1	158	2015	External	Ubiquiti Networks Inc.	101 - 1000	Professional	Social Engineering	Ubiquiti Networks Inc., the San Jose based man...
2	88	2013	External	University of Mississippi Medical Center	1,001 - 10,000	Healthcare	Lost and Stolen Assets	The University of Mississippi Medical Center i...
3	94	2013	Internal	Wells Fargo	Over 100,000	Finance	Privilege Misuse	former fire captain accessed the account of a ...
4	177	2016	External	Platte County Assessor Office	11 - 100	Public	Social Engineering	transferred money on spoofed email from CEO
...	...	...	...	...	...	...	...	...
115	2	2010	External	Scottrade	Unknown	Finance	Basic Web Application Attacks	Hackers gained unauthorized access to trading ...
116	120	2014	External	The Home Depot	Over 100,000	Retail	System Intrusion	The Home Depot, the world's largest home impro...
117	72	2013	Internal	Jumbo Chinese Buffet	Unknown	Accomodation	Everything Else	A man who waited tables at a Hazle Township re...
118	155	2015	External	Scottrade	1,001 - 10,000	Finance	Basic Web Application Attacks	Law enforcement officials contacted Scottrade,...
119	187	2017	External	Southern Oregon University	101 - 1000	Educational	Social Engineering	Business Email Compromise

120 rows × 8 columns

Figure 2: Illustration of dataset used for predictive cyber insurance analysis.

Figure 3 visualizes the dataset after the process of one-hot encoding and concatenation has been applied. One-hot encoding is a technique used to convert categorical variables (like industry names) into numerical format that machine learning models can understand. In addition, it also shown how the categorical columns have been transformed into binary columns, where each category becomes a binary feature. Concatenation refers to combining these encoded features with the rest of the dataset.

	ID	org_size_1 to 10	org_size_1,001 - 10,000	org_size_10,001 - 25,000	org_size_101 - 1000	org_size_11 - 100	org_size_25,001 - 50,000	org_size_50,001 - 100,000	org_size_Large	org_size_Over 100,000
0	137	0	0	0	0	1	0	0	0	0
1	158	0	0	0	1	0	0	0	0	0
2	88	0	1	0	0	0	0	0	0	0
3	94	0	0	0	0	0	0	0	0	1
4	177	0	0	0	0	1	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...
115	2	0	0	0	0	0	0	0	0	0
116	120	0	0	0	0	0	0	0	0	1
117	72	0	0	0	0	0	0	0	0	0
118	155	0	1	0	0	0	0	0	0	0
119	187	0	0	0	1	0	0	0	0	0

120 rows × 12 columns

org_size_Small	org_size_Unknown
0	0
0	0
0	0
0	0
0	0
...	...
0	1
0	0
0	1
0	0
0	0

Figure 3: Encoded and concatenated data after performing one-hot coding.

Figure 4 is the creation of a new data frame using the Natural Language Toolkit (NLTK) modules. NLTK is a library in Python used for natural language processing tasks. This shows the outcome of given dataset after going through the steps like tokenization (breaking text into words or tokens), stemming (reducing words to their root form), and possibly other preprocessing steps applied to the summary column. The resulting new data frame demonstrate how text-based information has been transformed into a structured format suitable for analysis.

	Actor	victim	victim.industry.name	summary
0	External	american tooling center	manufacturing	federal appeal court sided commercial policyho...
1	External	ubiquiti network inc	professional	ubiquiti network inc san jose based manufactur...
2	External	university mississippi medical center	healthcare	university mississippi medical center jackson ...
3	Internal	well fargo	finance	former fire captain accessed account well farg...
4	External	platte county assessor office	public	transferred money spoofed email ceo
...	...	...	...	...
115	External	scottrade	finance	hacker gained unauthorized access trading acco...
116	External	home depot	retail	home depot world largest home improvement reta...
117	Internal	jumbo chinese buffet	accomodation	man waited table hazle township restaurant han...
118	External	scottrade	finance	law enforcement official contacted scottrade u...
119	External	southern oregon university	educational	business email compromise

120 rows × 4 columns

Figure 4: Illustration of new data frame created using NLTK modules.

Figure 5 demonstrate the confusion matrix, which is a visualization commonly used in machine learning for evaluating the performance of classification models. It illustrated how this matrix is generated based on the predictions made by LR model, allowing for an assessment of how well the model's predictions align with the actual outcomes.

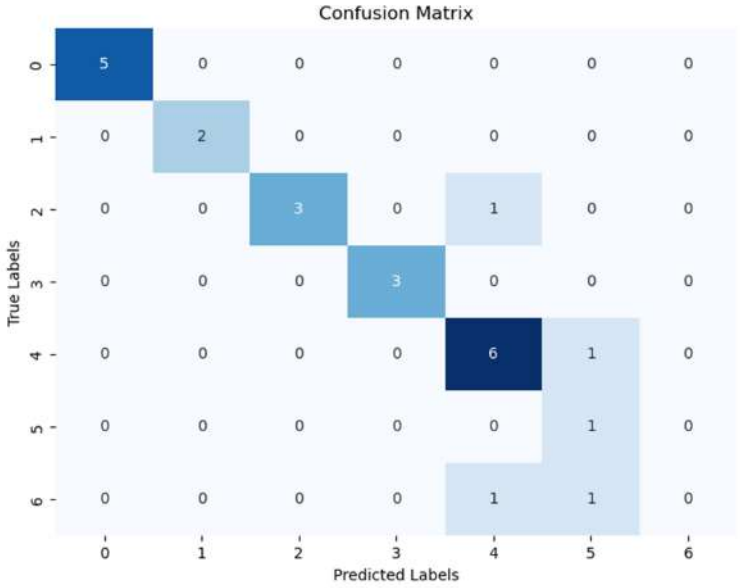

Figure 5: Confusion matrix obtained using Logistic Regression model.

Figure 6 displays the classification report obtained using LR model. A classification report is another evaluation tool for classification models. This figure showcases a detailed report that includes metrics such as precision, recall, and F1-score, for each class. These metrics help to assess the model's performance on individual classes and overall.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	5
1	1.00	1.00	1.00	2
2	1.00	0.75	0.86	4
3	1.00	1.00	1.00	3
4	0.75	0.86	0.80	7
5	0.33	1.00	0.50	1
6	0.00	0.00	0.00	2
accuracy			0.83	24
macro avg	0.73	0.80	0.74	24
weighted avg	0.82	0.83	0.81	24

Figure 6: Display of classification report obtained using Logistic Regression model.

Figure 7 presents a confusion matrix of GNB model, that helps to evaluate the performance of the proposed GNB model in classifying instances into different categories. Figure 7 displays a classification report specifically generated from the predictions of a GNB model. It provides metrics like precision, recall, and F1-score for each class, offering a comprehensive understanding of how well the GNB model performs in classifying different categories.

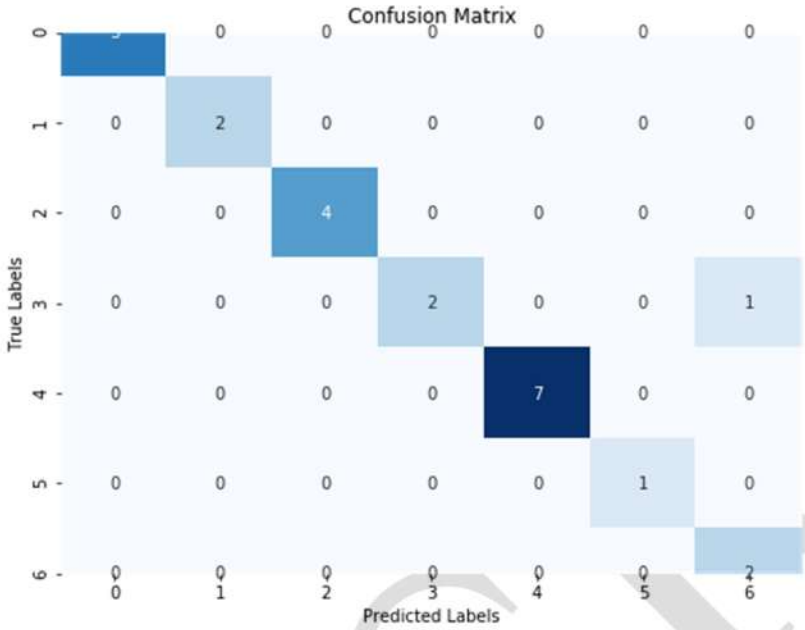

Figure 7: Confusion matrix obtained using gaussian Naïve Bayes model.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	5
1	1.00	1.00	1.00	2
2	1.00	1.00	1.00	4
3	1.00	0.67	0.80	3
4	1.00	1.00	1.00	7
5	1.00	1.00	1.00	1
6	0.67	1.00	0.80	2
accuracy			0.96	24
macro avg	0.95	0.95	0.94	24
weighted avg	0.97	0.96	0.96	24

Figure 8: Classification report obtained using gaussian Naïve Bayes model.

Table 1 presents a comprehensive performance comparison of two machine learning models, namely the LR model and the GNB model. The table evaluates the models based on several key metrics, including Accuracy, Precision, Recall, and F1-score. Each metric provides valuable insights into how well the models are performing in terms of correctly classifying instances and balancing various aspects of classification accuracy. From Table 1, the LR model achieves an Accuracy of 0.8333, indicating that around 83.33% of the instances in the dataset are correctly classified by the model. The Precision value for the LR model is 0.82, indicating that when the model predicts a certain class, it is accurate around 82% of the time. The Recall value for the LR model is 0.83, suggesting that the model successfully identifies around 83% of instances that actually belong to a specific class. The F1-score for the LR model is 0.81, which represents a balance between Precision and Recall, providing an overall measure of the model's accuracy that considers both false positives and false negatives.

On the other hand, the GNB model demonstrates a higher level of performance compared to the LR model. It achieves an impressive Accuracy of 0.9583, indicating that approximately 95.83% of instances are correctly classified. The Precision value for the GNB model is 0.97, reflecting a high level of accuracy when making

predictions for a specific class. The Recall value for the GNB model is 0.96, indicating that the model effectively identifies around 96% of instances that are actually part of a particular class. The F1-score for the GNB model is also 0.96, reaffirming the model's capability to balance Precision and Recall and achieve an overall accurate classification.

Table 1. Performance comparison of ML models in enhancing cyber security management with cyber insurance policy analysis.

Model	Accuracy	Precision	Recall	F1-score
LR model	0.8333	0.82	0.83	0.81
GNB model	0.9583	0.97	0.96	0.96

In summary, Table 1 highlights the performance characteristics of both the LR model and the GNB model. The GNB model outperforms the LR model in terms of Accuracy, Precision, Recall, and F1-score. This performance comparison showcases the strengths of the GNB model in correctly classifying instances and maintaining a balance between accurate positive predictions and the identification of actual positive instances. The obtained metrics provide valuable information for selecting an appropriate model for the predictive cyber insurance policy analysis, with the GNB model demonstrating superior performance based on the evaluation criteria presented.

5. CONCLUSIONS

The results of our study demonstrate the effectiveness of the proposed approach in predicting cyber insurance policy patterns. By using TF-IDF for feature extraction and the multinomial naive Bayes classifier for classification, we achieved high accuracy in predicting the patterns. This indicates that the combination of these techniques can be a valuable tool for insurance companies in understanding and predicting policy trends in the cyber insurance domain.

REFERENCES

[1] Kure, H.I., Islam, S. & Mouratidis, H. An integrated cyber security risk management framework and risk predication for the critical infrastructure protection. Neural Comput & Applic 34, 15241–15271 (2022). https://doi.org/10.1007/s00521-022-06959-2

[2] Albasheer H, Md Siraj M, Mubarakali A, Elsier Tayfour O, Salih S, Hamdan M, Khan S, Zainal A, Kamarudeen S. Cyber-Attack Prediction Based on Network Intrusion Detection Systems for Alert Correlation Techniques: A Survey. Sensors. 2022; 22(4):1494. https://doi.org/10.3390/s22041494

[3] Tsohou, A., Diamantopoulou, V., Gritzalis, S. et al. Cyber insurance: state of the art, trends and future directions. Int. J. Inf. Secur. 22, 737–748 (2023). https://doi.org/10.1007/s10207-023-00660-8

[4] J. Zhao, M. Shao, H. Wang, X. Yu, B. Li, X. Liu, Cyber threat prediction using dynamic heterogeneous graph learning, Knowledge-Based Systems, Volume 240, 2022, 108086, ISSN 0950-7051,https://doi.org/10.1016/j.knosys.2021.108086.

[5] M. Husák, J. Komárková, E. Bou-Harb and P. Čeleda, "Survey of Attack Projection, Prediction, and Forecasting in Cyber Security," in IEEE Communications Surveys & Tutorials, vol. 21, no. 1, pp. 640-660, Firstquarter 2019, doi: 10.1109/COMST.2018.2871866.

[6] Singh, A., Akhilesh, K.B. (2020). The Insurance Industry—Cyber Security in the Hyper-Connected Age. In: Akhilesh, K., Möller, D. (eds) Smart Technologies. Springer, Singapore. https://doi.org/10.1007/978-981-13-7139-4_16

[7] Hwang S-Y, Shin D-J, Kim J-J. Systematic Review on Identification and Prediction of Deep Learning-Based Cyber Security Technology and Convergence Fields. Symmetry. 2022; 14(4):683. https://doi.org/10.3390/sym14040683

[8] Sarker IH, Abushark YB, Alsolami F, Khan AI. IntruDTree: A Machine Learning Based Cyber Security Intrusion Detection Model. Symmetry. 2020; 12(5):754. https://doi.org/10.3390/sym12050754

[9] Lu, H., Zhang, G., Shen, Y. (2020). Cyber Security Situation Prediction Model Based on GWO-SVM. In: Barolli, L., Xhafa, F., Hussain, O. (eds) Innovative Mobile and Internet Services in Ubiquitous Computing . IMIS 2019. Advances in Intelligent Systems and Computing, vol 994. Springer, Cham. https://doi.org/10.1007/978-3-030-22263-5_16