Computers, Materials & Continua
DOI: 10.32604/cmc.2022.030235
Article

Tech Science Press

AI-Enabled Grouping Bridgehead to Secure Penetration Topics of Metaverse

Woo Hyun Park[1], Isma Farah Siddiqui[3] and Nawab Muhammad Faseeh Qureshi[2,*]

[1]Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, 16419, Korea
[2]Department of Computer Education, Sungkyunkwan University, Seoul, 03063, Korea
[3]Department of Software Engineering, Mehran University of Engineering & Technology, Jamshoro, Pakistan
*Corresponding Author: Nawab Muhammad Faseeh Qureshi. Email: faseeh@skku.edu
Received: 22 March 2022; Accepted: 17 May 2022

Abstract: With the advent of the big data era, security issues in the context of artificial intelligence (AI) and data analysis are attracting research attention. In the metaverse, which will become a virtual asset in the future, users' communication, movement with characters, text elements, etc., are required to integrate the real and virtual. However, they can be exposed to threats. Particularly, various hacker threats exist. For example, users' assets are exposed through notices and mail alerts regularly sent to users by operators. In the future, hacker threats will increase mainly due to naturally anonymous texts. Therefore, it is necessary to use the natural language processing technology of artificial intelligence, especially term frequency-inverse document frequency, word2vec, gated recurrent unit, recurrent neural network, and long-short term memory. Additionally, several application versions are used. Currently, research on tasks and performance for algorithm application is underway. We propose a grouping algorithm that focuses on securing various bridgehead strategies to secure topics for security and safety within the metaverse. The algorithm comprises three modules: extracting topics from attacks, managing dimensions, and performing grouping. Consequently, we create 24 topic-based models. Assuming normal and spam mail attacks to verify our algorithm, the accuracy of the previous application version was increased by ∼0.4%–1.5%.

Keywords: Metaverse; security; computational linguistics; grouping bridgehead; AI

1 Introduction

The governance basis of the metaverse is security. Anomaly detection is mainly employed for user safety protection in programs and applications. As such, companies mainly use natural language processing (NLP) technology to analyze a subject and character's movement intentions in texts exchanged between users. For example, the malicious intent of hackers through spam is a type of attack. Considerable research has been conducted on extracting textual information from datasets to analyze real spam. Machine learning (ML) methods, such as support vector machine (SVM) and logistic regression (LR), have been widely used as representative analysis methods to detect malicious

spam. Additionally, deep learning (DL) methods, such as recurrent neural networks (RNN) and deep neural networks, are widely used at present. However, as big data are generated within the metaverse where numerous users will come together in the future, security issues will become more prominent. Because the actions of users exchanged with various contents are recorded and affect others, a stable security strategy is required to analyze a subject. Various computational techniques, including term frequency and inverse document frequency (TF-IDF), word2vec, and RNN, have been developed, with accuracies of ∼95% or more. Particularly, RNN can now be derived with an accuracy of ∼98%. However, research is needed to increase accuracy for security to face various situations. In this study, we propose a grouping bridgehead (GB) algorithm for better performance and a more stable securing strategy than existing algorithms. The GB algorithm has three modules: extracting topics from attacks, managing dimensions, and performing grouping. It is a learning method that works on top of each other. In terms of F1-score, recall, and precision, we confirmed that the proposed algorithm performed highly compared with existing ML and DL counterpart algorithms. The novelties of this study are as follows.

- Improvement compared with existing performance through GB algorithm strategy for security.
- Comparative analysis through various topic-based models via NLP technology.
- A framework for security within the metaverse.

The remainder of this article is organized as follows. Section 2 examines cases from previous studies on how artificial intelligence (AI) and NLP technology are used in various fields; related studies on various data are also investigated. Section 3 explains the purpose of this study as well as the principles and development motives of the models proposed herein for security; further, it explains the methodology in order. Section 4 reports the collection process of data used in this study and the performance our methodology achieves. Section 5 summarizes this study, lists application areas of this study, and introduces future development directions.

2 Background and Related Studies of Metaverse

In this section, we present relevant studies within the metaverse.

It presented the characteristics of the metaverse and necessary technologies and problems for existing devices [1]. Particularly, in the metaverse, social problems such as ethics, calculability, and protection were described. Additionally, they explained three characteristics—sociality, multiplicity, and space–time relationship—and discussed convergence and the technical necessity for immersion as elements to actualize the metaverse. It performed an identity management inquiry for access management in a virtual metaverse and applied an automatic agent function [2]. Consequently, the Vortex platform was created; the performance of the Vortex Platform and compared with Vircadia and Sansar platforms in eight aspects (virtual world design, communication, avatar control, avatar design, scalability, content generation, support tool, etc.). From the evaluation, the Vortex platform scored slightly lower than other platforms, but it had strengths in assistive tools for e-learning, such as learning presentations. Although the content is relatively small, this is believed to be helpful for future educational platforms. It introduced new machine learning (ML) and deep learning (DL) techniques [3]. AI is employed in the era of big data to help decision-making in the medical field, which is emphasized as a technology that fuses the real world in the metaverse era of three-dimensional (3D) virtual space. It proposed a framework for healthcare within the metaverse [4]. First, keywords were extracted, searched, and the medical field was applied within the metaverse. The framework was analyzed using a database and applied to virtual reality (VR)/augmented reality. It helps to visualize 3D medical improvement. For example, there are three major diseases, disabilities, and exercise that

have been applied with wearable tools. Computing technology, medical technology, VR technology, network technology, etc. were selected as the technologies needed to develop the metaverse. However, it has been suggested that the issues that must be addressed before the technology implementation are particularly significant in healthcare, notably in terms of personal security and privacy. This is because a wide range of big data is threatened, including attacks by hackers on the system. It proposed a three-tier—infrastructure, interaction, and ecosystem—architecture for a vision for games, economy, and human interaction applicable in the metaverse [5]. For the metaverse development, it helped the campus map-professor and student classes using smartphones for ease of access for users and implementation of various functions.

3 Grouping Bridgehead

3.1 Motivation

There have been various techniques using machine learning, deep learning, and computational linguistic techniques to classify texts as well as observation records and articles used in healthcare data. It proposed a machine learning technique, using machine and deep learning techniques (long short term memory (LSTM), support vector machine (SVM), naive Bayes, logistic regression (LR), and k-nearest neighbor (KNN)), to collect comments on Reddit for approximately three months related to COVID-19 subjects (approximately 90), and performed sentiment analysis (positive, negative, and neutral) [6]. The study considered the length of repeated words, collection of negative words, complex meaning of words, processing of commercial words, emoticons that express adverbs of emphasis, and emoticons that express emotions. The LSTM model achieved a high accuracy of 81.15%. Further, the LR achieved an accuracy of 78.82%. This research will help health authorities understand diseases affecting people and the related strategies to overcome them. It investigated cases using NLP technology in relation to healthcare field documentation, and analyzed the cases using structural equations [7]. It reported the SMM4H [8]. To classify health information composed of syntactic identifiers for drugs on social media, three tasks (drug side effects classification, drug mention report classification, and normalization of expression of drug side effects) were performed. To perform the first task, sentiment analysis, n-gram, semantic function, cluster, machine learning (SVM and ensemble), and deep learning (DNN) techniques were used. A data imbalance problem for drug mentions was encountered, where different teams used under sampling and oversampling. In task 2, CNN, ensemble, inverse document frequency, and singular value decomposition were used. In task 3, LR, RNN, LSTM, and mean-value-based ensemble were used for regularization to correct spelling errors. Consequently, accuracies of 88.5% in task 2% and 88.7% with ensemble in task 3 were achieved. It solved the disease classification problem to some extent using machine learning methods (KNN, SVM, random forest (RF), extra trees, gradient boosting trees, extremely gradient boost, and LR) in the electronic health record (EHR) of ischemic stroke (IS) patients [9]. To employ unstructured data efficiently, features were selected, and the dimensions were reduced. Cross-validation was performed for overfitting. The evaluation results showed when XGBoost used both radiographic reports and progress records, the kappa value was 0.57, and cardiac embolism achieved the highest kappa value (0.63). Natural language processing (NLP) can be used to analyze epidemiological studies of disease classification. It attempted to analyze and extract sub-information (statistical information according to target group, number of patients, etc.) for disease classification in PubMed and EMBASE through NLP [10]. To extract only relevant data, 27 documents were manually selected. There were 12 symptom categories included in all papers (Articles, Emotional State, Circulatory and Respiratory Systems, Digestive System and Abdomen, Cognition and Perception, Pain, Fatigue and Sleep Disturbance, Nervous and Musculoskeletal, General, Skin and Subcutaneous Tissue, Urinary, and Symptoms Not

Reported). For classification, previously developed tools were used. This study is believed to be helpful for a systematic review. In this paper, the data were collected and pre-processed using the DATA automated crawling technique without manual filtering. It identified suicidal behaviors using electrical health records (HER) to prevent adolescent suicidal behaviors in advance [11]. A machine learning-based NLP method (random forest (RF)) was used as the identification method. First, a link was hypothesized between family and suicide attempts. With the consent of 73 adolescents in a psychiatric hospital, the term was constructed with clinical notes of atypical data related to psychotropic drugs. There was an average of approximately 130 notes per patient. The CUI was strengthened through literature research, and all character strings in the memo were converted to CUI. The Gini coefficient was used to measure impurities in depression and related terms. Twenty percent of the data were used for cross-validation. Thereafter, an ensemble classification technique was used to classify past suicide history. They found that psychotropic drugs contributed to suicide. The accuracy and area under curve (AUC) were 0.47 and 0.68, respectively. Previous NLP studies conducted in cyberspace are as follows.

A multilingual learning environment was built within the metaverse to collect information, and attempts were made to analyze records through problem-based learning [12]. Specifically, students' activities were monitored by collecting places, gaze directions, and chat contacts. The authors calculated statistical data for chatting, divided the relationship between score index and readability into seven steps, tested about eight readability algorithms such as automated readability index, gunning fog score and so on, and linked them with grades. Specifically, it helped analyze text complexity and give advice at the right time to the students who were discussing it. To run a metaverse environment related to aircraft maintenance training on an independent device such as smart glasses, It used a convolutional neural network (CNN) model to recognize Korean and English mixed voice speech and classify commands such as voices and gestures [13]. As a result of a mixed evaluation of recall and precision with 2,000 audio samples and ∼1,100 Korean and English data points, command speech recognition worked at 95.7% and language prediction achieved 99.6%. It is believed that the independent device can guide aviation maintenance trainees as guidelines. It studied the impact of user experience-based design to reflect the real world in the metaverse [14]. As a result of questionnaire analysis and survey, attractiveness, usability, and interaction were positive, proving the relevance of user experience-based in terms of metaverse identification and immersion. It is believed that these studies will help advance the metaverse and virtual reality (VR) in the future. Additionally, for the metaverse development, we anticipated that user security problems will emerge in the future metaverse era, and accordingly, we tried solving the user security problems. It defined the metaverse as a space for computer art, such as movies, products, music, poetry, and novels, and then investigated various socially influential factors [15]. They highlighted the security issue posed by digital informatization for the elements of these creations to be integrated into the metaverse. Accordingly, interdisciplinary research involving computer-linked technologies is required. In this study, we tried solving the security problem in the metaverse using AI technology. It has been studied in weight allocation and optimization techniques to solve computational cost and complexity in vehicle re-identification [16]. For this, the channel importance was considered using CNN and SE block, and a method was used to further supplement the features by adding an intermediate layer. In addition, learning was carried out in consideration of image occlusion. Experimental results showed that it is higher than the widely known classification of vehicle identification. It performed multi-label learning for feature analysis in Chinese [17]. As a result of analyzing the existing algorithms, the performance of SVM has been proven to be effective.

3.2 Information Protection with Subject and Deep Learning

In this study, we focused on the topics of deep learning (DL) and machine learning (ML) and documentation for our research. It created a system for monitoring widespread penetration to counter network attacks on social media [18]. They tried to detect by applying the NLP model. Specifically, the federated neural network and LDA models were used. They used LR, unigram, big gram, n-gram, etc. in the model and selected the largest subject in 20 iterative sampling until convergence to capture a specific signal in a distributed denial-of-service (DDoS) attack. This was performed by infiltrating 50 attacks against a corpus of DDoS manually collected for a network attack experiment. As a result of the experiment, our algorithm outperformed comparison methods in terms of F1 score and improved by 4% and 25% compared with a neural network and trend approach, respectively. It is believed that the analysis result will provide the user with a stopping point. It conducted a study on how the classification performance of DL varies according to the distribution of the CIFAR10 dataset and visualized common characteristics [19].

3.3 Overall Structure of Security in Natural Language Processing Technology

We address security issues in the meta space in this study. First, an anonymous attack through malicious email is assumed. The device (cell phone, computer, etc.) that received it becomes infected with a virus. Types of damage from attacks include falsifying invoices, falsifying resumes of job seekers, and passing on information to others. Additionally, files containing vital information can be spread to an unspecified number of people. This is mainly done as an attack on URL links. Email is an unstructured text, and security has been studied using NLP technology for email attacks. It studied information management technology to secure competitiveness for security by applying NLP to information processing outsourcing [20]. Information security risk classification was divided, for example, attackers that cause confusion, such as falsification when exchanging documents, changing meaning by confusing synonyms, negative words, and affirmative words, changing a sentence as if it was approved by a human, making changes of the content contrary to a contract, and denying data agent access. Some issues also interfere with the model, causing NLP to malfunction, such as the modification of the data provider and damage by attacking inside the black-box model. As an experiment, a survey was conducted with nine participants in a scenario. This is believed to help companies in their strategic exploration for information management. It also conducted a study on information management strategies by performing interview improvement scenarios from participants belonging to various organizations by grafting NLP for Information technology outsourcing (ITO) [21]. There are attacks via hacker alarms in the metaspace.(See Fig. 1).

The purpose and type of damage are to plant a virus, which causes severe damages such as monitoring and manipulating the screens of users' devices (interphones, laptops, etc.) to shoot and distribute illegal images. This can occur primarily to target media connected to a home network. Security has been studied using NLP technology in preparation for attacks on Internet Protocol (IP) and domains of home networks. It studied sequence search and cost need for datasets [22]. Maliciousness was identified by extracting domain and IP address relationships from log files and network sequence traffic volumes using NLP. They updated data on the basis of facts and classified relevant characteristics as malicious entities through clustering using distance. The evaluation was performed using different broadband network datasets of 2,000 people collected over a week in Europe, and the number of malicious products using related viruses was reported. As IP and domain increase, the bed vector tends to increase significantly. Additionally, hackers in metacyber space attack through messages. For the purpose and type of damage of messenger attacks, they can maliciously

access users' information through voice phishing, delete information, perform forced termination, interfere with users, and even give personal insults to users.

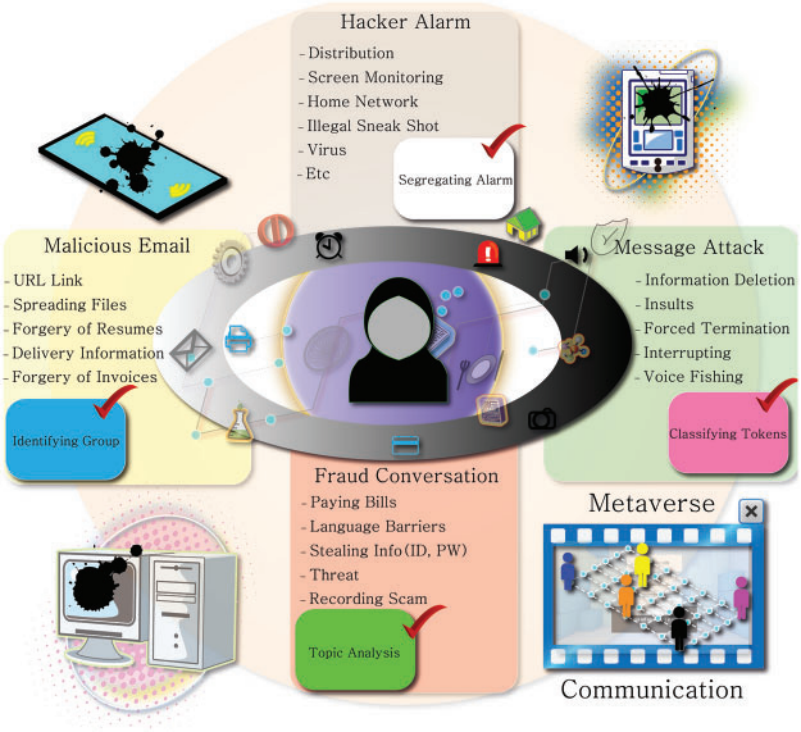

Figure 1: The chart of analysis for security in metaverse communication

Conversations may also occur with fraudulent intent. The hacker's purpose is to implement language barriers to generate bills, steal users' login information, perform recording scams, and then threaten them. NLP technology is used to classify unstructured texts on social network services (SNS). It proposed speech-and word-based metrics to solve the attack problem of private backdoors [23]. The attack success rate, detection success rate, accuracy of the backdoor model, and concealability of the backdoors were considered for evaluation. Product review and SNS data were used for sentiment analysis and toxicity detection. As a result of the evaluation, all contaminated data showed an accuracy of 94% or more, and RW had the highest detection rate. Concealment was highest in sentence level and low in Stealthy Backdoor Attack with Stable Activation. The backdoor had a high attack probability. This will be helpful in the method for security as an NLP study to remove the harmfulness of backdoors on SNS. Another method using language computing is the bridgehead strategy. The task flow is shown in Fig. 2. The components for a bridgehead strategy are as follows.

- Unique ID/password, etc. (pilot)
- Avatar and Behavior (customed)
- Cyber environment (platform)
- Controller (If needed)
- Storage service (cloud, etc.)
- Software/hardware environment (spec)
- User communication (produced by contents in metaverse and pilot)
- Agent tech

- Devices (PC, notebook, etc.)
- Computing NLP
- Etc.

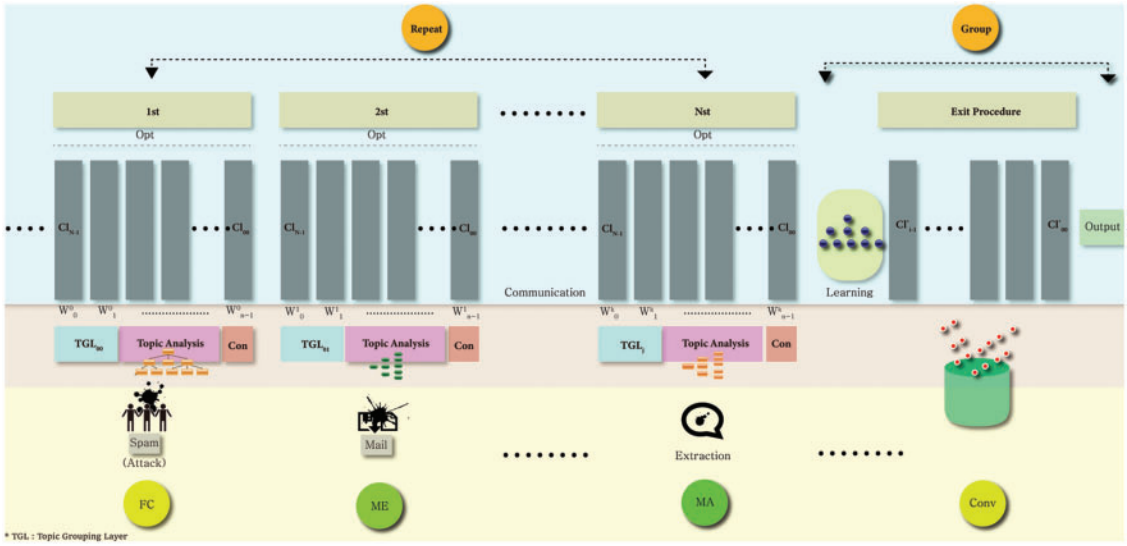

Figure 2: Architecture of grouping bridgehead to secure penetration topics for security in metaverse

Multiple learning methods are used for high-quality learning data and algorithm performance. It analyzes the topic according to the fraud spam attack, classifies topics with high similarity, and converts it to a weight calculation technique for multiclassification. Then, optimizations are performed. Afterward, the topic is analyzed according to the malicious mail attack, and similar topics are classified and weighted for classification. Then, the dimension is managed and the optimal value is converted. Repeatedly, for message extraction, the topic is divided among subtopics, multi classified by calculating weights, and then dimensional decomposition is performed. As such, the optimal value is converted up to N times, and the obtained values are continuously combined. After iteration, until the final procedure is reached, group learning is performed, and the final output is extracted. The VR and NLP technology interaction systems excelled in maintenance technology, which will be discussed in detail in Section 4. The method algorithm in this study has four main proposed modules. The first module classifies the subject, which uses the latent Dijkstra allocation algorithm. In the past, and the formula for the study are as follows [24]:

$$TTIS(\mathrm{t}, \mathrm{d}, \mathrm{T}, \mathrm{D}, n, r) \tag{1}$$

$$TIS(t, d, T, D, n, r) \tag{2}$$

$$TFTIDF(t, d, T, D) \tag{3}$$

Term frequency-topic inverse document frequency with a singular value decomposition (TTIS) is a model that combines the dimensional decomposition technique with regularization to strengthen the sparsity problem generated by matrix of document and term by analyzing the document topic as $\mathrm{Basis}_{\_i}$. The advantage of TTIS is that it creates a synergistic effect over the conventional language methods on social media. Second, term frequency-topic inverse document frequency (TFTIDF)

performed standardization and topic-based document analysis on specific documents, inverse numbers, etc. using a probabilistic technique. As a result, it was found to be more effective than the conventional model. Term frequency inverse document in singular value decomposition (TIS) calculated the importance from the existing word frequency document frequency and performed singular value decomposition.

$\sigma_i(t)$ is a model that performs malicious detection classification by LR. A linear vector is generated for the existing document-word topic classification weights {w0, w1, w2, ..., w_{n-1}}. For data processing, we trained several computational models to detect malicious classifications. Among them were SVM and k-nearest neighbor (KNN). SVM classified the variables as Wx+b in a hyperplane equation, whereas KNN grouped the variables by calculating the distance over a set of t∈T. A typical expression for KNN is as follows [25]:

$$d_i = \sqrt{\sum_{i=1}^{p} (x_{2i} - x_{1i})^2} \tag{4}$$

Representatively, the ensemble model was employed. Multiple trees were configured to classify properties; malicious attacks were classified. The expression is as follows [26]:

$$fi_i = \frac{\sum_{j:node\ j\ splits\ on\ feature\ i} ni_j}{\sum_{k \in all\ nodes} ni_k} \tag{5}$$

After calculating the predicted distribution for the basis classifier in a post processing method for each procedure, the layers were increased, and optimization was performed.

$$\begin{cases} Opt_r\ (Model_i) \\ \quad Cl_M \end{cases} \tag{6}$$

When the layers are continuously stacked, there is a role of merging them. The topic grouping layer is allocated in a stack structure to improve the existing topic and to end the procedure.

The expression is as follows:

$$\text{concat} \begin{cases} p\,(spam|TGL)\,,\ if\ N < v \\ \quad Conv\ Stacking \\ \quad Procedure\ exit \end{cases} \tag{7}$$

where N denotes the number of processes, and v denotes variable cases.

Topic group layer (TGL) calculates the probability of spam given the j_{th} category in collection_v(fraud conversation, email, alarm, message attack, otherwise) to maximize the likelihood and merges the topic model. The expression is as follows:

$$\text{TGL} = \rho\left(\text{spam}|\text{Category of Cl}_j\right) \tag{8}$$

The final E_n^*(DL, ML) results by recruiting DL and ML complex data processors for each case for each bridgehead N_{th} stack. The expression of Estimator I-1 among all is as follows:

$$\text{E} = \text{Learning}\left(\text{N}_{st}^*\,, \text{Grouping}\right)^* \tag{9}$$

For the entire dataset and prediction data (X′, Y′), true positive (TP), true negative (TN), false positive (FP), and false negative (FN), respectively, representing true positive, true negative, false positive, and false negative are evaluated to measure accuracy after constructing a classification matrix. The model formula is as follows [25]:

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{FP} + \text{FN} + \text{TN})} * 100\% \tag{10}$$

Algorithm 1: Pseudo Code of the Group Bridge Algorithm in Meta-Verse

Input: Original Corpus
Initialization:
  1. training Data S
  2. parameters $\{\text{N}_{layer}, B_{cl}, W, T_j, P_n, r\}$ *in learning*
Procedure:
for $c \leftarrow 0$ **to** $N - 1$ **do**
  Stor_{id} *in MC*
  for $k \leftarrow 0$ **to** $v - 1$ **do**
    Preprocesed
    Parsing & Sorting & Build V_{col}
    Build
    $\text{T}_k() \leftarrow$ *using LDA*
    $\text{Vec}() \leftarrow$ *using Eq.* (1)
    Selective Dimensional Management_r
    $\text{TGL} \leftarrow$ *using Eq.* (6)
    for $j \leftarrow 0$ **to** B **do**
      $\text{Cl}_j \leftarrow$ *using Eq.* (4)
      $\text{Proc}_N \leftarrow$ *using Eq.* (7)
      Learning
      Optimization
      Until Convergence
    end for
    Calculation using Estimator
    Checking procedure and storage
    Concat
    Repeat :
    Updating W_i
    $\text{T}_k() \leftarrow$ *using Eq.* (1)
    Generate Vec()
    Selective Dimensional Management_r
    $\text{TGL} \leftarrow$ *using Eq.* (6)
    $\text{Cl}_j \leftarrow E_{N-1}$
    $\text{Grouping}_i()$
    Learning
    Procedure Exit
  end for
end for

Precision was measured for basic classification detection. In the proposed algorithm, when FP and TP are given, the calculation method is to divide TP by FP + TP as a ratio. The malicious classification is detected by measuring recall. Meanwhile, given FN and TP, the calculation method is to divide TP by FN + TP.

Equity was also calculated by measuring the F1 scores.

The description of the algorithmic process of the grouping bridgehead (GB) model is as follows. There are parameters for estimating the binary dataset and bridgehead grouping learning assuming that the malicious data set have penetrated. N represents the number of layers, B_{cl} represents the number of classifiers, W is the topic weight, T_j is topic count, P_n is the estimator's procedure, and r is the dimension split count. The metaverse stores user communication data in storage and processes necessary for analysis. After parsing the collection in MC, you will build it. By constructing the topic vector, the overall topic classification vector required for bridgehead analysis will be constructed. Optional dimension management for computing on big data. After configuring a data processor using ML and DL, calculations for each procedure are learned. This is repeated until convergence. After updating w according to the collection, we compute the topic probabilities, generate a vector, and compute the final model iteratively. After calculating the maximum stacking length according to the collection, the endpoint is determined and grouped. This GB strategy will excel in maintaining the interaction between VR and NLP technology.

4 Technology Analysis and Experimental Results

4.1 Baseline

Tab. 1 summarizes the environment and data in which the methodology was tested; performance was evaluated in terms of accuracy (ACC), F1-score (F1), precision (P), and recall (R). Downloaded datasets and existing methods in the literature [24,27–36] were used for performance evaluation assuming an anonymous attack through malicious email. As a result of classification, liked and separated corpus were distinguished. The corpus was divided into training and test data for learning and evaluation, respectively. The learning and normalization rates were set to 0.001 and 0.01, respectively.

Table 1: Baseline

Dataset	Environment	Algorithms
UCI Mail dataset	Tensorflow/Keras/Scikit-learn/Python 3.9 Windows 10-home 64bit	Grouping Bridgiehead/ Existing Models

4.2 Classification Performance

The performance measured for each layer of the basic models employed in this study is summarized. As the bridgehead layer increased, the performance was compared in various manners. The accuracy measurements of the training data are shown in Fig. 3. In the training data, the model of Basis 2 had the best accuracy of 1. Bases 1, 3, and 4's accuracies were 0.9993, 0.9996, and 0.9939, respectively. In general, all values were greater or equal to 0.99, indicating good performance. The gated recurrent unit (GRU) model had an accuracy of 0.9932, which was lower than the Basis step. The convolutional neural network (CNN) model had an accuracy of 0.992, which was also lower than the Basis level. The layer is believed to cause overfitting with respect to the critical point. The accuracy measurements of the test data are shown in Fig. 4. In the test data, Basis 3 yielded the highest accuracy (0.98134). Bases 1, 2, and 4 had accuracies of 0.9803, 0.9792, and 0.9738, respectively. In general, all values were 0.97 or higher, which resulted in good performance compared with the malignancy detection of existing models.

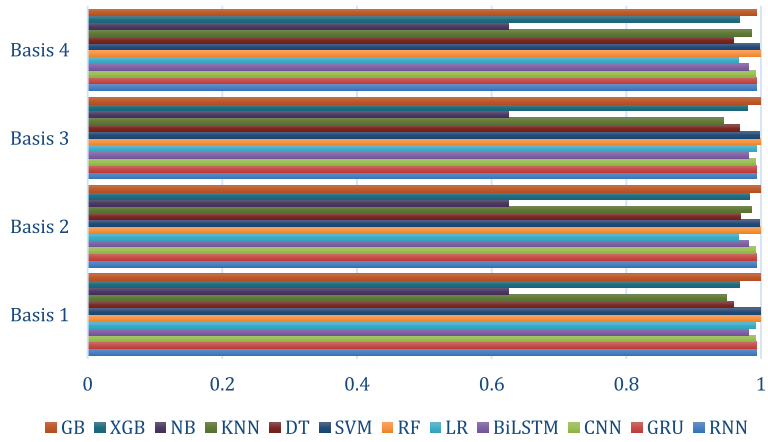

Figure 3: The accuracy of malicious email attack classification in each layer (training)

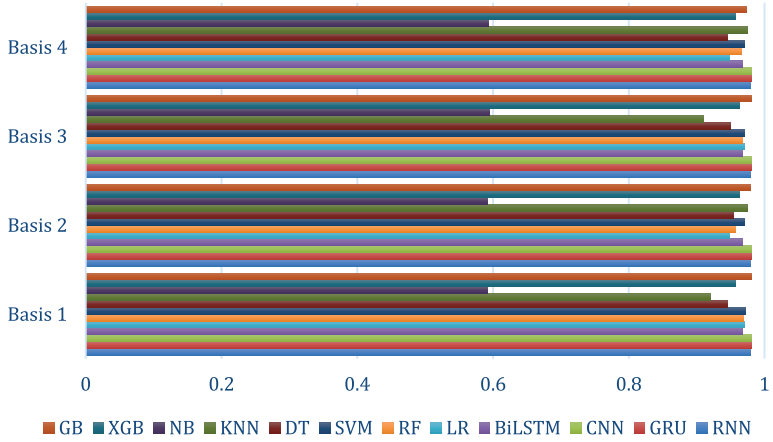

Figure 4: The accuracy of malicious email attack classification in each layer (test)

Typically, GRU had an accuracy of 0.9816 due to five times training, and recurrent neural network (RNN), k nearest neighbor (KNN), support vector machine (SVM), and logistic regression (LR) had accuracies of 0.9797 (Basis 4), 0.9749, 0.9723 (Basis 1), and 0.9705 (Basis 1), respectively. The proposed model has room for better performance than by the adjustment of GB strategy parameters. Obviously, there is an optimal number of layers, and a huge synergy is expected when using the GB strategy.

4.3 Precision in Each Layer

Fig. 5 shows the precision measured for each layer of the basic models employed in this study. A performance comparison was conducted as the bridgehead layer was increased. As a result of measuring the precision index of the test data, Basis 1 had the highest precision (1). Bases 2, 3, and 4 all achieved 0.99. LR, random forest (RF), SVM, and KNN (Basis 4) all achieved 0.99. Bidirectional long-short term memory (BiLSTM) and RNN achieved 1 and 0.98, respectively. NB showed relatively low precision (0.05). In general, the subject-based GB strategy yielded the best precision of 0.99 or higher in all steps.

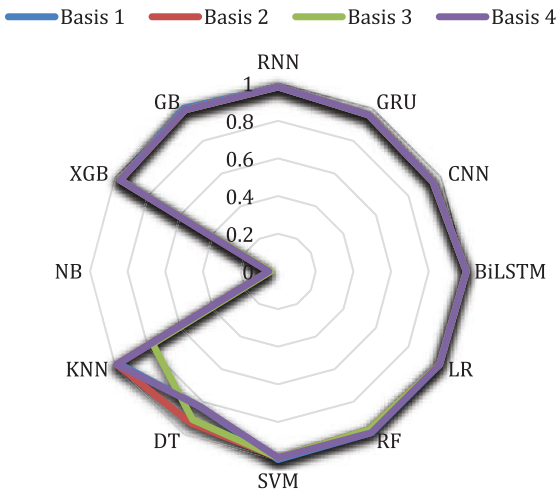

Figure 5: The precision of malicious email attack in each layer

4.4 Recall in Each Layer

Fig. 6 shows the results of filtering for malicious group identification and malicious subject tokens and classifications from hacker attacks. As the bridgehead layer increased, a performance comparison was conducted on the recall index of the test data. As a result, Basis 3 showed the highest recall value of 0.87. Bases 1 and 2 had 0.85, and Basis 4 had 0.8. CNN achieved the highest score of 0.91; GRU, RNN, and KNN had recall values of 0.9, 0.87, and 0.82, respectively. In general, the theme-based GB strategy achieved a recall value of 0.8 or higher in all stages.

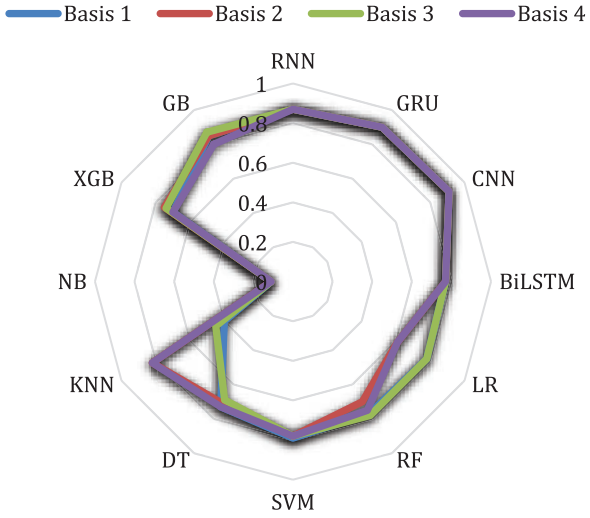

Figure 6: The recall of malicious email attack in each layer

4.5 F1 in Each Layer

Fig. 7 shows the harmonic average of the precision and recall for the malicious group identification and malicious subject tokens from hacker attacks for each layer of the basic models employed in this

study. As the bridgehead layer increased with respect to class imbalance, a performance comparison was conducted on the F1 score of the test data. As a result of measuring the F1 index of the test data, Bases 1 and 3 had the highest statistical value of 0.92. It was shown to be the best model. Bases 2 and 4 achieved 0.91 and 0.89, respectively. GRU and CNN had a value of 0.92, RNN was estimated to be 0.91, SVM was 0.88, and RF and LR were estimated to be 0.87. NB performed the worst at all layers; its performance was even worse than that of RM. This indicates that there is no synergistic effect in the security malicious detection bridgehead technique, suggesting a zero frequency for irregular and continuous indicators in relation to prior probability. In general, the theme-based GB strategy had an F1 score of ∼0.9 or higher in all stages. Based on previous studies, LR performed well, but another surprising finding related to security malware detection while conducting this study is that SVM and RF were very good. Because SVM had little effect on erroneous data, it could yield performance improvement and is interpreted as being able to prevent overfitting to sparse data. RF was interpreted as having an effective advantage in processing missing values and large volumes of big data. This topic-based GB strategy will be effective for technology processing big data in the metaverse.

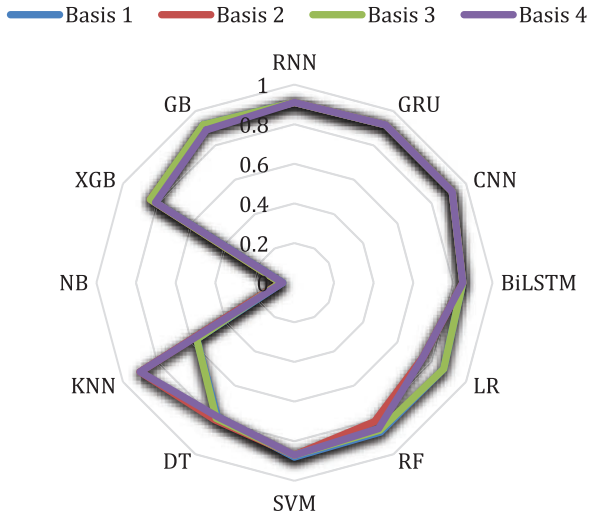

Figure 7: The F1 of malicious email attack in each layer

5 Conclusion

In this study, we applied AI-based linguistic computing technology to effectively process big data for security in cyberspace in view of the metaverse era, which comprises several assets. Accordingly, a topic-based grouping bridgehead model was developed to solve the security response problem. Each model presented in this study comprises three modules. There are topic classification, establishment, and confirmation of malicious beachhead groupings as well as dimension management. We found that the filtering performed well for the token classification strategy. This shows the effect of reinforcing the feature problem as a result of previous studies. As a result of the experiment, the accuracy, F1 score, recall, and precision improved, respectively, by ∼1%–30%, 1%–17%, 1%–18%, and 1%–22% or more compared with the existing models. The proposed methodology proved to be effective in protecting users from malicious infiltration of hackers within the twin space. As a result of conducting a security strategy study to prevent penetration from hacker attacks, we found that filtering performed well for topic classification, malicious group identification, dimension management, and token classification

strategies. The proposed model has the potential to be further developed by extending it to attacks through email, messages, and alarms, and communication attacks that users may experience in the metaverse. The interaction of experts is required to develop the model and will assist in decision-making.

Acknowledgement: This research was supported by the SungKyunKwan University and the BK21 FOUR(Graduate School Innovation) funded by the Ministry of Education(MOE, Korea) and National Research Foundation of Korea(NRF)

Funding Statement: This work was supported by the BK21 FOUR Project. W.H.P received the grant.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] H. Ning, H. Wang, Y. Lin, W. Wang, S. Dhelim *et al.,* "A survey on metaverse: The state-of-the-art, technologies, applications, and challenges," arXiv preprint arXiv:2111.09673, 2021.

[2] A. Jovanović and A. Milosavljević. "VoRtex metaverse platform for gamified collaborative learning," *Electronics*, vol. 11, no. 3, pp. 317–337, 2022.

[3] K. Aggarwal, M. M. Mijwil and A. H. Al-Mistarehi, "Has the future started? The current growth of artificial intelligence, machine learning, and deep learning," *Iraqi Journal for Computer Science and Mathematics*, vol. 3, no. 1, pp. 115–123, 2022.

[4] D. Chen and R. Zhang. "Exploring research trends of emerging technologies in health metaverse: A bibliometric analysis," Available at SSRN 3998068, 2022.

[5] H. Duan, J. Li, S. Fan, Z. Lin, X. Wu *et al.,* "Metaverse for social good: A university campus prototype," in *Proc. of the 29th ACM Int. Conf. on Multimedia*, NY, pp. 153–161, 2021.

[6] H. Jelodar, Y. Wang and R. Orji. "Deep sentiment classification and topic discovery on novel coronavirus or COVID-19 online discussions: NLP using LSTM recurrent neural network approach," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 10, pp. 2733–2742, 2020.

[7] D. Ionescu, "Deep learning algorithms and big health care data in clinical natural language processing," *Linguistic and Philosophical Investigations*, vol. 19, pp. 86–92, 2020.

[8] A. Sarker, M. Belousov and J. Friedrichs. "Data and systems for medication-related text classification and concept normalization from twitter: Insights from the social media mining for health (SMM4H)-2017 shared task," *Journal of the American Medical Informatics Association*, vol. 25, no. 10, pp. 1274–1283, 2018.

[9] R. Garg, E. Oh, A. Naidech and K. Kording. "Automating ischemic stroke subtype classification using machine learning and natural language processing," *Journal of Stroke and Cerebrovascular Diseases*, vol. 28, no. 7, pp. 2045–2051, 2019.

[10] T. A. Koleck, C. Dreisbach and P. E. Bourne. "Natural language processing of symptoms documented in free-text narratives of electronic health records: A systematic review," *Journal of the American Medical Informatics Association*, vol. 26, no. 4, pp. 364–379, 2019.

[11] N. J. Carson, B. Mullin, M. J. Sanchez, F. Lu and K. Yang. "Identification of suicidal behavior among psychiatrically hospitalized adolescents using natural language processing and machine learning of electronic health records," *Plos one*, vol. 14, no. 2, e0211116, 2019.

[12] K. T. Nakahira, N. R. Rodrigo and R. Taguchi. "Design of a multilinguistic problem based learning learning environment in the metaverse," in *2nd Int. Symposium on Aware Computing*, Tainan, pp. 298–303, 2010.

[13] A. Siyaev and G. S. Jo. "Towards aircraft maintenance metaverse using speech interactions with virtual objects in mixed reality," *Sensors*, vol. 21, no. 6, pp. 2066–2087, 2021.

[14] J. Jeon, "The effects of user experience-based design innovativeness on user-metaverse platform channel relationships in South Korea," *Journal of Distribution Science*, vol. 19, no. 11, pp. 81–90, 2021.
[15] L. H. Lee, Z. Lin, R. Hu, Z. Gong, A. Kumar *et al.,* "When creators meet the metaverse: A survey on computational arts," arXiv preprint arXiv:2111.13486, 2021.
[16] X. R. Zhang, X. Chen, W. Sun and X. Z. He, "Vehicle Re-identification model based on optimized densenet121 with joint loss," *Computers, Materials & Continua*, vol. 67, no. 3, pp. 3933–3948, 2021.
[17] J. He, C. Wang, H. Wu, L. Yan and C. Lu, "Multi-label Chinese comments categorization: Comparison of multi-label learning algorithms," *Journal of New Media*, vol. 1, no. 2, pp. 51–61, 2019.
[18] N. Chambers, B. Fry and J. McMasters. "Detecting denial-of-service attacks from social media text: Applying nlp to computer security," in *Proc. of the 2018 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Louisiana, vol. 1, pp. 1626–1635, 2018.
[19] S. T. Kim and H. G. Kim. "Understanding which images degrade deep learning classification performance," *TECHART: Journal of Arts and Imaging Science*, vol. 8, no. 4, pp. 14–18, 2021.
[20] B. M. Bhatti, S. Mubarak and S. Nagalingam. "Information security implications of using NLP in IT outsourcing: A diffusion of innovation theory perspective," *Automated Software Engineering*, vol. 28, no. 2, pp. 1–29, 2021.
[21] B. M. Bhatti and S. Mubarak. "NLP-Based enhancement of information security in ITO-A diffusion of innovation theory perspective," in *2020 35th IEEE/ACM Int. Conf. on Automated Software Engineering Workshops (ASEW)*, Australia, pp. 112–117, 2020.
[22] G. Siracusano, M. Trevisan, R. Gonzalez and R. Bifulco, "Poster: On the application of NLP to discover relationships between malicious network entities," in *Proc. of the 2019 ACM SIGSAC Conf. on Computer and Communications Security*, NY, pp. 2641–2643, 2019.
[23] W. Yang, Y. Lin, P. Li, J. Zhou and X. Sun, "Rethinking stealthiness of backdoor attack against nlp models," in *Proc. of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th Int. Joint Conf. on Natural Language Processing*, online, vol. 1, pp. 5543–5555, 2021.
[24] W. Park, D. R. Shin and N. M. F. Qureshi. "Pseudo NLP joint spam classification technique for big data cluster," *Computers, Materials & Continua(CMC)*, vol. 71, no. 1, pp. 517–535, 2022.
[25] D. A. Anggoro and N. D. Kurnia. "Comparison of accuracy level of support vector machine (SVM) and K-nearest neighbors (KNN) algorithms in predicting heart disease," *World Academy of Research in Science and Engineering*, vol. 8, no. 5, 2020.
[26] F. M. J. M. Shamrat, Z. Tasnim, P. Ghosh, A. Majumder and M. D. Z. Hasan, "Personalization of job circular announcement to applicants using decision tree classification algorithm," in *2020 IEEE Int. Conf. for Innovation in Technology (INOCON)*, Bangluru, pp. 1–5, 2020.
[27] W. Park, D. R. Shin and N. M. F. Qureshi. "Effective emotion recognition technique in NLP task over nonlinear Big data cluster," *Wireless Communications and Mobile Computing*, vol. 2021, 2021.
[28] S. Sartaj and A. F. Mollah. "An intelligent system for spam message detection," *Intelligent Systems. Springer*, pp. 387–395, 2021.
[29] G. Sousa, D. C. G. Pedronette and J. P. Papa. "SMS spam detection through skip-gram embeddings and shallow networks," in *Findings of the Association for Computational Linguistics: ACL-IJCNLP*, PA, pp. 4193–4201, 2021.
[30] V. Vishagini and A. K. Rajan. "An improved spam detection method with weighted support vector machine," in *2018 Int. Conf. on Data Science and Engineering (ICDSE)*, Kochi, pp. 1–5, 2018.
[31] G. Chetty, H. Bui and M. White. "Deep learning based spam detection system," in *2019 Int. Conf. on Machine Learning and Data Engineering (iCMLDE)*, Taipei, pp. 91–96, 2019.
[32] S. Nandhini and J. M. KS. "Performance evaluation of machine learning algorithms for email spam detection," in *2020 Int. Conf. on Emerging Trends in Information Technology and Engineering (IC-ETITE)*, Vellore, pp. 1–4, 2020.

[33] G. Jain, M. Sharma and B. Agarwal. "Spam detection on social media using semantic convolutional neural network," *International Journal of Knowledge Discovery in Bioinformatics (IJKDB)*, vol. 8, no. 1, pp. 12–26, 2018.

[34] S. S. Ali and J. Maqsood. "Net library for SMS spam detection using machine learning: A cross platform solution," in *2018 15th Int. Bhurban Conf. on Applied Sciences and Technology (IBCAST)*, Islamabad, pp. 470–476, 2018.

[35] A. Ghourabi, M. A. Mahmood and Q. M. Alzubi. "A hybrid CNN-LSTM model for SMS spam detection in arabic and English messages," *Future Internet*, vol. 12, no. 9, pp. 1147–1154, 2020.

[36] Y. Zhang, P. F. Liu and J. T. Yao. "Three-way email spam filtering with game-theoretic rough sets," in *2019 Int. Conference on Computing, Networking and Communications (ICNC)*, Canada, pp. 552–556, 2019.