International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-11 Issue-3, January 2022

Protecting Against Malicious Code Injection in Reviews on Web Applications

Nivetha. V, Shalini. S, Deepa. R

Abstract: ***Malicious code injection is done by the attackers or hackers mainly in reviews, these are the fake user identities that are created by the attackers, through sending continuous links to some user till the user clicks on that link. When user clicks the link, that particular users identity will be stolen by the attackers which is done like the phishing, generally these reviews which are maliciously injected are mixed with the original users review of the product in a website. To identify these malicious injection attacked reviews, in this paper the Naive Bayes Classifier (NBC) algorithm is used. Then to eliminate the unwanted disturbed data in the website the Natural Language Processing technique is used. The Natural Language Processing (NLP) technique, which removes unwanted data by understanding the text or words in review. Then the length of the reviews in the given sample dataset is reduced for easy understanding by using the Principal Component Analysis (PCA) algorithm which reduces the dimensionality of the reviews. Then user input review is compared with the sample dataset and classified as good review and bad review and also detected as malicious or not by using the Naïve Bayes Classifier algorithm which is used for the classification of the objects.***

Keywords: Malicious, Naïve Bayes Classifier, Natural Language Processing, Principal Component Analysis

I. INTRODUCTION

The machine learning is defined as the one in which the machine is trained and tested. The machine learning has two phases as the training phase and the testing phase. There are three types of machine learning as Supervised learning, Unsupervised learning and Reinforcement learning. The Supervised learning is defined as the sample dataset is given to the machine by the user for training. The Unsupervised Learning is defined as the machine itself considers the dataset on its own. The Reinforcement Learning is defined as the one in which it is a machine learning training method based on rewarding desired behaviors and punishing undesired ones. Malicious code injection is defined as the one in which it occurs when an attacker exploits an input validation flaw in system to inject malicious code.

Manuscript received on December 11, 2021.
Revised Manuscript received on January 17, 2022.
Manuscript published on January 30, 2022.
* Correspondence Author
V. Nivetha*, P.G Graduate, Department of Computer Science Engineering, Prince Dr. K. .Vasudevan College of Engineering and Technology, Ponmar, Chennai, India, E-mail: vnivetha72@gmail.com
S. Shalini, Assistant Professor, Department of Computer Science Engineering, Prince Dr. K.. Vasudevan College of Engineering and Technology, Ponmar, Chennai, India, E-mail: s.shalinicse@princedrkvasudevan.com
R. Deepa, M.E, Head of Department of Computer Science Engineering, Prince Dr. K.. Vasudevan College of Engineering and Technology, Ponmar, Chennai, India, E-mail: rkdeepa14@gmail.com

There are some types of malicious code injection attacking types, the two of them are Co-visitation injection attack and Profile injection attack. The Co-visitation injection attack is defined as the attack in which the attackers or hackers keep on viewing the website but without buying a product, the attackers will give a review for product in a website. The Profile injection attack is defined as the attack in which the attackers or hackers give a review for a product in a website through fake user identity. This proposed paper, is based on the Machine Learning algorithms where the sample dataset is trained to the machine and then tested for seeking the result or the output. The need for the protection of the people or the users from trusting the malicious reviews given for the product in the web applications, is to detect and eliminate those malicious reviews injected. These malicious reviews are the malicious code injected reviews which are used for the attacking purpose in reviews by the attackers. These attackers or the hackers create those malicious code injected reviews. The malicious code injected reviews are injected by the attackers or the hackers when the website service is low in the web applications. Hence these malicious injection attacked reviews are detected and eliminated by using the Natural Language Processing (NLP) technique. This technique can be applied for the elimination of the unwanted disturbed data's in the web applications. Naïve Bayes Classifier (NBC) algorithm can be implemented for the classification of the objects. Here this particular algorithm is applied for the classification of the given user input reviews into good reviews and bad reviews. This algorithm is one of the Supervised Machine Learning algorithms. This algorithm applied is very fast in accuracy and can be used to make the real time predictions easily. For easy understanding of the reviews which are in the sample dataset the Principal Component Analysis (PCA) algorithm is used. This PCA algorithm is often used as the dimensionality-reduction technique. Here this algorithm is implemented for reducing the length of the reviews in the sample dataset. And also this algorithm is one of the Machine Learning algorithms.

Malicious reviews detection and elimination mainly aims to protect the people or the users from trusting the malicious code injected reviews and also the user input review is classified into good review and bad review. There are some types of malicious injection attacks, the two of them are Co-visitation injection attack and Profile injection attack. The pre-processing aims at the elimination of the disturbed data which are the malicious injection attacked reviews. By enhancing the profile injection behaviors and co-visitation injection behaviors the disturbed data is eliminated.

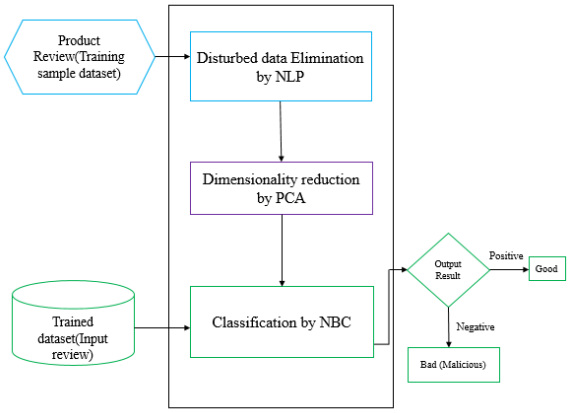
And also the sentiment identification of the reviews in the sample dataset is done, which identifies the reviews emotions as whether the review sentiment is kind of positive or negative or malicious. The central aim of the PCA (Principal Component Analysis) is the dimensionality-reduction of the reviews in the sample dataset. In which the length of the reviews are reduced into a easier format for easy understanding. The advantage of PCA is that it improves the visualization of the data and has high performance speed and hence works fast. And the constraint number is also fixed according to the sample dataset given to the machine for training. The main focus for the NBC (Naïve Bayes Classifier) is to classify the objects. In which here, the user input reviews are compared with the sample dataset given for training the machine and then classified as positive(good) review and negative(bad) review and also detected as malicious or not. It is also compared with the five other algorithms to calculate the accuracy value of the algorithm. The advantage of NBC is that it has the highest accuracy value.

II. EXISTING SYSTEM

The existing system presents a divide and conquer strategy to detect two types of injection attacks as Profile Injection Attacks(PIA) and Co-visitation Injection Attacks(CIA) for online recommender systems. First the disturbed data is eliminated by enhancing the two injection attacks called as PIA and CIA using the List Then Eliminating(LTE) technique. Then the dimensionality is reduced that is the length of the data is reduced for the better understanding by using the Gradient Descent Algorithm. The IMIA-HCRF technique is defined as the Identify Malicious Injection Attacks using Higher Order Conditional Random Fields.Finally, the classification is done and the result is given as positive and negative by using the IMIA-HCRF technique. The Accuracy rate value of the IMIA-HCRF technique is about from 0.6 to point 0.7.

III. PROPOSED SYSTEM

The main reason for this project is that the people or the users when buying a product in the website, they mainly see the reviews for their reference and for knowing about the product. In that case when they see the reviews they unknowingly considers the malicious reviews also as the reviews given by the original users and they judge and come to a decision about the product, since these malicious reviews are mixed with the original users reviews. These malicious reviews are the malicious code injected reviews by the attackers through the malicious injection attacks when the website service is low. Malicious code injection is defined as occurring when an attacker exploits an input validation flaw in system to inject malicious code. There are some types of malicious code injection attacking types, the two of them as follows. The Co-visitation injection attack is defined as the attackers or hackers keep on viewing the website but without buying a product, the attackers will give a review for product in a website. The Profile injection attack is defined as the attackers or hackers give a review for a product in a website through fake user identity. Hence to overcome this problem, this project is proposed. The main goal of this project is to save the users or the people from trusting these malicious reviews. By using Naïve Bayes Classifier (NBC) algorithm which has highest accuracy, the reviews are identified as good or bad and also detected as malicious or not.



In the proposed system, the sentiment data's that is the unwanted disturbed data's are first eliminated from the given dataset by using the Natural Language Processing (NLP) technique. Secondly the constraint number is fixed according to the given sample, hence the dimensionality of the given dataset is reduced and the important data alone is extracted using the Principal Component Analysis (PCA). Finally, the input is given which is compared with the sample dataset and by using the Naive Bayes Classifier (NBC), it is classified into fake reviews and original reviews in which the original reviews are further classified as positive (good) and negative (bad) which is displayed with the algorithm accuracy percentage value. The Eliminating Noise Data in Website (ENDW) technique eliminates the noisy data with a greater efficiency. The Natural Language Processing (NLP) technique has higher rate of sentiment identification. The Principal Component Analysis (PCA) algorithm improves the visualization of the data and has high performance speed and hence works fast. NBC algorithm is defined as the one which is an algorithm that is used for the classification of the objects is called as NBC algorithm. The Naïve Bayes Classifier (NBC) algorithm is the algorithm that has the higher algorithm accuracy value.

A. Pre-processing

In preprocessing stage, first the training of the machine is done. This comes under the training phase in machine learning. This machine training is done by collecting the dataset from any of the web applications such as Amazon, extras. And this dataset for training the machine which is taken from the data source called as the kaggle. Then the enhancement of Profile injection behaviours and Co-visitation injection behaviours is done. By enhancing the profile injection behaviors and co-visitation injection behaviors the disturbed data is eliminated. The malicious code injected reviews are injected by the attackers or the hackers when the website service is low in the web applications.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Malicious code injection is defined as the one in which it occurs when an attacker exploits an input validation flaw in system to inject malicious code. There are some types of malicious code injection attacking types, the two of them are Co-visitation injection attack and Profile injection attack. There are fake user identities that are created by the attackers, through sending continuous links to some user till the user clicks on that link. When user clicks the link, that particular users identity will be stolen by the attackers which is done like the phishing. Generally these reviews which are maliciously injected are mixed with the original users review of the product in a website.The disturbed data is eliminated by using the Natural Language Processing (NLP) technique. Natural Language Processing has applications in various sectors like tourism, where a tourist can get information about the famous tourist spots in a particular city, the hotels available, best restaurants nearby etc. Another important application of NLP is chatbot (Chat Robot) that can be used for voice or textual interactions. The goal is a computer capable of "understanding" the contents of documents, including the contextual nuances of the language within them. The technology can then accurately extract information and insights contained in the documents as well as categorize and organize the documents themselves. NLP reduces the gap between computer and human by combining the database query into meaningful forms. Natural Language Processing (NLP) technique is defined as the one in which it is used to eliminate the disturbed data in the websites. In simple words, the steps involved in this elimination process using NLP technique are as follows, First the meaning of given input reviews from the dataset are analyzed for better machine understanding. Then the arrangement of the review words from the dataset are done. The disturbed data like symbols, meaningless words, extra are separated in the input dataset given. Then finally, those separated data's are alone removed or eliminated from the input dataset given.

Fig1. Sample dataset

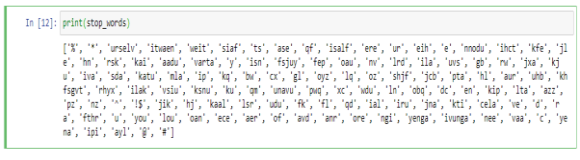


Fig2. Eliminating disturbed data

B. Feature Extraction

In feature extraction stage, first the constraint number is fixed according to the given sample training dataset in the preprocessing stage. Hence after the removal of disturbed data in the preprocessing stage, the remaining dataset is given as the input to this feature extraction stage. Then the dimensionality of the given dataset is reduced that is the length of each review in the dataset is minimized for easy understanding. Hence the important data alone is extracted using the Principal Component Analysis (PCA) algorithm. The Principal Component Analysis (PCA) is often used as the dimensionality-reduction technique.Principal component analysis (PCA) is the process of computing the principal components and using them to perform a change of basis on the data, sometimes using only the first few principal components and ignoring the rest. PCA is used in exploratory data analysis and for making predictive models. PCA is defined as an orthogonal linear transformation that transforms the data to a new coordinate system such that the greatest variance by some scalar projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on.In simple words, the steps involved in this dimensionality-reduction process using Principal Component Analysis (PCA) algorithm are as follows, First the dataset which is the output from the preprocessing module is collected. The data headings from the dataset are first listed for machine reference. Then the unwanted data heading are alone spotted separately. Finally, those unwanted data heading are then removed to reduce the length of the dataset for machine easy understanding. The Principal Component Analysis (PCA) algorithm improves the visualization of the data. The PCA algorithm also has the high performance speed. And hence the PCA algorithm works fast. The Principal Component Analysis improves the performance of the Machine Learning algorithm. The Principal Component Analysis results in the high variance and thus improves the visualization. PCA algorithm is one of the Machine Learning algorithms.

Fig3. Before dimensionality-reduction process

Fig4. After dimensionality-reduction process

C. Classification

In classification stage, the review is given as the input for the classification of output. This input review which is given, is compared with the sample dataset which is the output of the feature extraction module. After the comparison is done, the input review is classified either as an positive (good) or as an negative (bad) and also detected as malicious or not by using the Naive Bayes Classifier (NBC) algorithm.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The Naïve Bayes Classifier algorithm is defined as the one which is an algorithm that is used for the classification of the objects is called as NBC algorithm.Naïve Bayes classifiers are highly scalable, requiring a number of parameters linear in the number of variables (features/predictors) in a learning problem. Naive Bayes is a simple technique for constructing classifiers and models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. There is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: all naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable. For example, a fruit may be considered to be an apple if it is red, round, and about 10 cm in diameter. NBC algorithm is one of the supervised Machine Learning algorithms.The independent feature model that is, the naïve Bayes probability model. The naïve Bayes classifier combines this model with a decision rule. One common rule is to pick the hypothesis that is most probable; this is known as the maximum a posteriori or MAP decision rule. In simple words, the Naïve Bayes Classifier(NBC) algorithm working procedure in steps are as follows, In train classifier, after giving the sample dataset into the machine, the machine forms the unique words from the dataset given. Machine then also forms the frequency of each word in a document from the unique words formed. Now after this, machine removes the unwanted words by referring from the dictionary of stop word. Then if negative label, then calculate the probabilities for negative. If positive label, then calculate the probabilities for positive. If malicious label, then calculate the probabilities for malicious. For the probability calculation of each class such as positive, negative and malicious : First taking the positive outcomes. Then computing the probability for each unique words in positive label by using the formula,

$$P(wk/+) = \frac{(nk+1)}{n} + |\text{Vocabulary}|$$

Where, nk : Number of times word k occurs in positive case.
n : Number of words in positive.

Vocabulary : Total unique words.

Repeating the step 4.1.2 for negative label and malicious label. While testing for unknown words, then nk = 0 is used and its probability is found for all positive, negative, malicious labels. In test classifier, the input test data is given into the machine. Now repeating the step 3 for the input test data given. Comparing each reviews in the input test data with the trained dataset labels, and separating the reviews into positive label, negative label and malicious label. Now probability for each reviews in all three labels are calculated. Now, adding all the review probabilities of each label to get the total values of each class. After adding, the classification is done now and the total value of each class is displayed, which is the final output result.The Naïve Bayes Classifier(NBC) algorithm is compared with the few other algorithms to calculate the accuracy value of the algorithm. The Naïve Bayes Classifier algorithm which is used in the proposed system has the higher algorithm accuracy value up to 0.9. The Naïve Bayes Classifier is simple and easy to implement and it handles both continuous and discrete data's. The NBC is very fast in accuracy and can be used to make real time predictions.

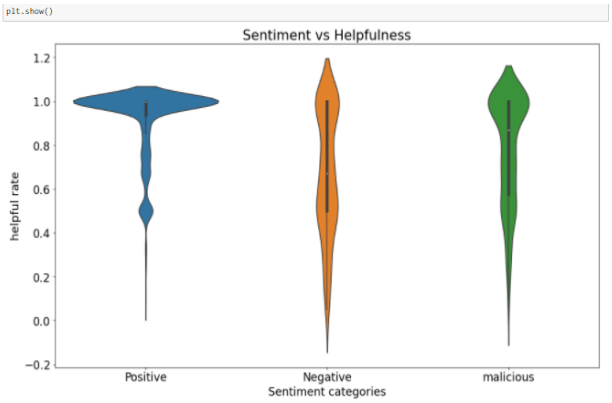


Fig5. Output of malicious reviews with positive and negative from trained dataset

IV. RESULT AND DISCUSSION

First and foremost the machine has been successfully trained. And then the malicious code injected reviews are been detected by using the Naïve Bayes Classifier algorithm. Then the disturbed data in the review are eliminated by using NLP technique. Mainly the affected malicious review is successfully detected by using the NBC algorithm. Hence the main goal of saving the people or users from unknowingly trusting the malicious code injected reviews has been successfully achieved. This NBC algorithm used for classification process for the detection of reviews as good or bad and also as malicious or not, has the higher algorithm accuracy detection rate value up to 0.9.

V. CONCLUSION AND FUTURE WORK

The Detection of malicious injection attack in reviews with algorithms high accuracy value is proposed in this paper. This proposed paper also detects the type of the review from the given input review whether it is positive(good) review or negative(bad) review and also detected as malicious or not. We train the machine in this paper by giving the sample datasets, hence the machine compares the input with the trained datasets and provide the output result. The algorithm used in this paper for the classification of output results as good or bad and is malicious or not, the accuracy value is calculated by comparing it with few algorithms to show our used algorithm in the proposed system has the highest accuracy value. In the future work, the further enhancement process of the project is that the malicious code injected Uniform Resource Locator which is the expansion of URL that is the Links of the web applications can be identified. And then can also be detected as whether the uniform resource locator of the web applications are malicious or not in the near future.

Retrieval Number: 100.1/ijitee.C97200111322
DOI: 10.35940/ijitee.C9720.0111322
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

ACKNOWLEDGEMENT

I would like to express great fullness to P.G. Department of Computer Science Engineering, Prince. Dr. K. Vasudevan College of Engineering and Technology, Ponmar.

REFERENCES

1. Naufal Riza Fatahillah, Pulut Suryati, Cosmas Haryaan,"Implementation Of Naive Bayes Classifier Algorithm On Social Media (Twitter) To The Teaching Of Indonesian Hate Speech", 2017
2. Revathy M, Minu Lalitha Madhavu, "Efficient Author Community Generation On Nlp Based Relevance Feature Detection", 2017
3. Yunjing An, Shutao Sun, Shujuan Wang, "Naive Bayes Classifiers for Music Emotion Classification Based on Lyrics", 2017
4. Ershad Sharifahmadian, Alireza Ahmadian, "Adaptive Signal Processing Algorithm for Remote Detection of Heart Rate (HR) Using Ultra-Wideband Waveforms based on Principal Component Analysis", 2009
5. Moch. Fadli Shadiqin Thirafi, Faisal Rahutomo, "Implementation of Naïve Bayes Classifier Algorithm to Categorize Indonesian Song Lyrics Based on Age", 2018
6. Rushrukh Rayan, Md. Sabir Hossain, and Asaduzzaman, "Compression of Large-Scale Image Dataset using Principal Component Analysis and K-means Clustering", 2019
7. Uma M, Sneha V, Sneha G, Bhuvana J, Bharathi B, "Formation of SQL from Natural Language Query using NLP", 2019
8. Sidharth Prasad Mishra, Uttam Sarkar, Subhash Taraphder, Sanjay Datta, Devi Prasanna Swain, Reshma Saikhom, Sasmita Panda and Menalsh Laishram, "Multivariate Statistical Data Analysis- Principal Component Analysis (PCA)", 2018
9. Dr. N. Srinivasan, Anandaraj Selvaraj, "Mobile Based Data Retrieval using RDF and NLP in an Efficient Approach", 2017
10. Anggit Dwi Hartanto, Ema Utami, Sumarni Adi, Harish Setyo Hudnanto, "Job Seeker Profile Classification of Twitter Data Using the Naïve Bayes Classifier Algorithm Based on the DISC Method", 2019

AUTHORS PROFILE

V. Nivetha, PG Graduate, Computer Science Engineering, Prince Dr. K. Vasudevan College of Engineering and Technology, Ponmar, Chennai, India, E-mail: vnivetha72@gmail.com

S. Shalini, Assistant Professor of Computer Science Engineering, Prince Dr. K. Vasudevan College of Engineering and Technology, Ponmar, Chennai, India, E-mail: s.shalinicse@princedrkvasudevan.com

R. Deepa, M.E, Head of the Department, Computer Science Engineering, Prince Dr. K.. Vasudevan College of Engineering and Technology, Ponmar, Chennai, India, E-mail**:** rkdeepa14@gmail.com