Most commonly used machine learning algorithms for cybersecurity incident reports classification

Veneta Yosifova, Antoniya Tasheva, Roumen Trifonov
Technical University of Sofia, Bulgaria,
venetay@tu-sofia.bg, atasheva@tu-sofia.bg, r_trifonov@tu-sofia.bg

***Abstract**: The challenge for information security industry is creating reliable models for classification of unified incident reports, so-called "tickets", as to eliminate the human factor as a cause of delay and mistakes. With the increasing amount of cybersecurity incidents, the automatic detection of anomalies and trends in incidents response systems is essential. Machine learning methods are used to speed up response and increase the quality in the management of incidents reports. Automatic classification of the tickets according to a common taxonomy allows computer security professionals to follow international standards for the next steps in processing the incidents. This paper makes a survey of the most popular and common used machine learning algorithms for cybersecurity incidents classification.*

***Keywords:** Machine Learning, Incident Response, Cybersecurity Incident Reports, Tickets, Automatic Classification*

1. INTRODUCTION

For information security industry it is important creating reliable models for classification of unified incident reports, so-called "tickets", in order to eliminate the human factor as a cause of delay and mistakes. In the practice, service desk operators, who first receive information about an incident, have to evaluate each case in order to assign appropriate values to the ticket's attributes. Most common attributes are type, urgency, category, team, and etc. After that each ticket is sent to next level of specialists who have to identify next steps for the incident processing. In the case of international companies, outsourced services or international cooperation, incident handling teams may be located in different parts of the world, and unification according to internationally recognized best practices and standards is a key part in incident response process.

Machine learning methods are used to speed up response and increase the quality in the management of incidents reports. Automatic classification of the tickets according to a common taxonomy allows computer security professionals to follow international standards for the next steps in processing the incidents.

In addition, tickets analysis makes it possible to "identify anomalies and trends, as well as detect unusual patterns in the operations. Such analysis is hard to do manually especially for large accounts with complex organization and scopes" [1]. For that reasons, it is appropriate to automate such tasks using artificial intelligence algorithms and machine learning methods. Similar approach is used in related works as "highlighting of suspicious activity, anomalies in extracted log files, volatile memory, or drive images" [2] for providing valuable clues about incidents. Identifying "various cyber incidents either previously seen or unseen, is a key issue to be solved urgently" [3]. In any of these case using machine learning methods for automatic classification of the incidents or incident reports is of great help for large organizations.

2. MACHINE LEARNING ALGORITHMS FOR CYBERSECURITY INCIDENTS CLASSIFICATION

There is no universally accepted definition of what a cybersecurity incident is, and each organization have to define what an "incident" is in their case, so that in event of an emergency, the incidents response teams to know how to respond [4]. One general definition is given by vendor independent best practice collection, formerly known as Information Technology Infrastructure Library – ITIL [5], where “an incident is an unplanned interruption to an IT service or reduction in the quality of an IT service" [6]. This is too broad definition but international classification such as ENISA Threats Taxonomy [7] allows computer security professionals to follow best standards for recognizing incidents and the next steps in their processing.

Currently the main approaches for solving machine learning tasks are supervised machine learning, unsupervised machine learning or semi-supervised learning. There are two types of tasks in the supervised machine learning - regression and classification tasks. Unsupervised machine learning solves clustering problems.

After analysis of the scientific literature focused on incident response issues, we can highlight most often used algorithms for classification based on supervised learning such as Support Vector Machine (SVM) and K-Nearest Neighbor (KNN). Naive Bayes classification and other methods based on regression problems such as the Decision Trees.

2.1. Term Frequency-Inverse Document Frequency (TF-IDF)

Many applications use TF-IDF as a text prepossessing and preparing the input for classification algorithms. On the basis of statistics TF-IDF [8] determines the frequency of occurrence of words in the ticket’s description text. This helps to extract the meaning from a text without having to make a complete semantic analysis of the ticket description. Term Frequency is the number of times a term occurs in a document. The weight of a term that occurs in a document is proportional to the term frequency [9]. Inverse Document Frequency decrease the weight of terms that occur very frequently in the document and increases the weight of terms that occur rarely. For instance, the word “the” is very common however not a good keyword to include in a text classification process.

2.2. Support Vector Machine (SVM)

The reason SVM to be preferred for such kind of problems is because this supervised learning method solves classification problems [10] and because this method is known to be good for text categorization tasks [11]. Usually Support Vector Machine method is used "to get the class for a given incident description" [12] based on text categorization which is essential for extracting the keywords from an incident report.

Statistic methods as TF-IDF and Boolean weighting for text representation produce smaller input vectors for SVM classifier.

As explained in [13] “incident description has an important role in the categorization of incidents. As shown training the data with only the incident description has an accuracy result of 86%”. But training the data only with nominal attributes the overall accuracy is 43%. Definitely classification with text description attribute gives much better results.

2.3. K-Nearest Neighbor (KNN)

This is a metric algorithm that measures distances between objects for automatic classification. The KNN algorithm uses the Euclidean distance as a classification indicator to identify the nearest neighbors [14]. It is suitable for solving classification problems and again can be combined with TF-IDF for text prepossessing.

On the same dataset from the example of the previous section, the KNN results are a little worse than SVM method. The accuracy with text description attribute is 80%. Nominal attributes have 42% accuracy. That's why one of the most important part of solving classification problems is finding relevant "features", characteristics that adequately reflect objective dependencies on the classification model [15].

2.4. Naive Bayes classification (NB)

Naive Bayes alongside Support vector machines are good text classification algorithms. NB classifier relies on probabilities of events and is based on Bayes Naive Theorem [16]:

$$P\langle A|B\rangle = \frac{P\langle B|A\rangle P(A)}{P(B)} \tag{1}$$

The probability for an incident's category is calculated for all categories. This method is good for text and document classification. NB classifier count the number of times each word occurs and ignore the ordering. Each "document can be represented as a p-vector of counts (a word frequency histogram)" [17]. It's used to find features in incident reports that belong to a specific category.

In [18] the authors use Naive Bayesian method for incident classification and their incident classification approach resulted in 70% accuracy with 1000 features.

2.5. Decision Trees

Another intuitive way to classify are Decision trees [19]. This method is type of supervised machine learning for regression problems. Decision tree can be used for both regression and classification problems. The "core concept of this model is a division of the dataset into smaller datasets which are conceptualized on the descriptive features. This division takes place until the smallest dataset is obtained containing the data points to be categorized under a single specific label" [20]. The trees consist of nodes with leaves. At each node, a decision is made based on the value of one of the features by which it is classified. Going down the tree and reaching a leaf, the classification is complete because each leaf is associated with a class. Applying this approach can help in classifying the received tickets to a classification scheme, the different types of incidents are separate leaves from the decision tree.

In [13] are cited results where Decision trees method combined with TF-IDF have around of 90% accuracy. On the same dataset SVM produces similar accuracy but Naive Bayes presents different results with 85% for TF-IDF and even lower 55% with Term Frequency only.

2.6. Neural networks

Using Neural Network for classification tasks is another option. In [21] the authors use Softmax classifier based on regression Neural Network. The regression problem is generalized to classification problems where the class label can take more than two possible values. They compare the accuracy using only ticket subject and subject and description of the ticket. Again the results are better if ticked description is used ~83% versus ~80% for subject only. Next they compare Naive Bayes and NN. When the two categories were simultaneously chosen, the best overall accuracy achieved 85.8% for NB classifier.

In Table 1 are summarized most commonly used machine learning algorithms for cybersecurity incidents classification based on studied papers, sources and researches.

Tab. 1: Machine learning algorithms for cybersecurity incidents classification.

Classification method	**Method type – tasks**	**Best sides**
Support Vector Machine – SVM	Supervised ML – Classification	Suitable for text categorization
K-Nearest Neighbor - KNN	Supervised ML – Classification	Easy for implement and text categorization
Naive Bayes	Supervised ML – Classification	Good for text and document classification
Decision Trees	Supervised ML – Regression/Classification	Good for regression and classification problems
Neural networks	Supervised ML – Regression→Classification	Detect anomalies and classification problems

All of them are supervised machine learning methods suitable for solving classification problems.

3. CONCLUSION AND FUTURE WORK

We have made a review of the most commonly used machine learning algorithms for cybersecurity incidents classification. We presented results based on the reviewed papers but despite we cited accuracity for all these approaches it is not possible to compare all of them at once without having same base such equal dataset and dataset size, chosen features and categories.

In our future work we are going to train the methods on real incident tickets dataset. This should allow us to compare their real accuracy and performance.

4. ACKNOWLEDGEMENTS

The work presented in the paper is supported by research grant 202ПД0023-09 financed by the Research and Development Centre of the Technical University of Sofia, Bulgaria and the national science program “Information and Communication Technologies for Common Digital Market in Science, Education and Security” managed by Ministry of Education and Science

5. REFERENCES

[1] Li, Ta & Liu, Rong & Sukaviriya, Noi & Li, Ying & Yang, Jeaha & Sandin, Michael & Lee, Juhnyoung. (2014). Incident Ticket Analytics for IT Application

Management Services. Proceedings - 2014 IEEE International Conference on Services Computing, SCC 2014. 568-574.

[2] C. Nila, I. Apostol and V. Patriciu, "Machine learning approach to quick incident response," 2020 13th International Conference on Communications (COMM), Bucharest, Romania, 2020, pp. 291-296.

[3] Sarker, I.H., Kayes, A.S.M., Badsha, S. et al. Cybersecurity data science: an overview from machine learning perspective. J Big Data 7, 41 (2020).

[4] Luttgens, J.T., Pepe, M., & Mandia, K. (2014). Incident Response & Computer Forensics, 3rd Edition, p. 50.

[5] "ITIL's the name - you won't wear it out!", AXELOS, June 2018. Visited August 2020. https://www.axelos.com/news/blogs/june-2018/itils-the-name-you-wont-wear-it-out

[6] ITIL Service Operation. Ed.3. 2011, p. 99.

[7] Latest version of ENISA's Threat Taxonomy. Updated in September 2016. Last visited August 2020. https://www.enisa.europa.eu/topics/threat-risk-management/threats-and-trends/enisa-threat-landscape/threat-taxonomy/view

[8] B. Trstenjak, S. Mikac, and D. Donko, 'KNN with TF-IDF based framework for text categorization', Procedia Eng., vol. 69, pp. 1356–1364, 2014.

[9] Luhn, Hans Peter (1957). "A Statistical Approach to Mechanized Encoding and Searching of Literary Information" (PDF). IBM Journal of Research and Development. 1 (4): 309–317.

[10] Mark Stamp, Introduction to Machine Learning with Applications in Information Security-CRC (2018). p. 95.

[11] T. Joachims, 'Text categorization with support vector machines: Learning with many relevant features', Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 1398, pp. 137–142, 1998.

[12] R. Gupta, K. H. Prasad, and M. Mohania, 'Automating ITSM incident management process', 5th Int. Conf. Auton. Comput. ICAC 2008, vol. 1, pp. 141–150, 2008.

[13] Silva, S., Pereira, R. & Ribeiro, R. (2018). Machine learning in incident categorization automation. In 13th Iberian Conference on Information Systems and Technologies (CISTI). Cáceres: IEEE.

[14] Y. Song, J. Huang, D. Zhou, H. Zha, and C. L. Giles, 'IKNN: Informative K-Nearest Neighbor Pattern Classification', Proc. Eur. Conf. Princ. Pract. Knowl. Discov. Databases, pp. 248–264, 2007.

[15] Trifonov, Roumen, Manolov, Slavcho, Tsochev, Georgi, Pavlova, Galya. (2020). Automation of Cyber Security Incident Handling through Artificial Intelligence Methods. WSEAS TRANSACTIONS on COMPUTERS.

[16] Jaya Aiyappan, Naive Bayes Classifier for Text Classification, August 2019. Visited 2020. https://medium.com/analytics-vidhya/naive-bayes-classifier-for-text-classification-556fabaf252b

[17] Murphy, K.P. (2006). Naive Bayes Classifiers. Technical Report, October 2006.

[18] Gupta, R., Prasad, K. H., Luan, L., Rosu, D., & Ward, C. (2009). Multi-dimensional Knowledge Integration for Efficient Incident Management in a Services Cloud. IEEE International Conference on Services. 2009.

[19] Anita C. Faul, A Concise Introduction to Machine Learning-CRC Press (2020)., p. 123.

[20] Gupta, Brij, M. Sheng. - Machine learning for computer and cyber security principles, algorithms, and practices-CRC Press (2019), p.66.

[21] G. Son, V. Hazlewood, and G. D. Peterson, 'On Automating XSEDE User Ticket Classification', 2014.