IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

Efficient Bankruptcy Prediction using Machine Learning

Yuvaraj S.E[1] and A. Poongodi[2]
PG Student, Department of Computer Applications[1]
Professor, Department of Computer Applications[2]
Vels Institute of Science, Technology and Advanced Studies, Pallavarm, Chennai.
yuvarajelanthezhian2000@gmail.com and poongodimca1979@gmail.com

Abstract: *This study is based on a meta-analysis of bankruptcy prediction using machine learning. The data on these studies was collected on six levels: algorithms, data balance, variable categories, variables types, industry, and region. The aim of this project is to analyze the determinants of accuracy in bankruptcy prediction models. To achieve this aim, Mixed Effects models were developed. The results obtained show that while some factors are significant determinants for the accuracy of machine learning models in bankruptcy prediction (algorithm, data balance, industry, region), some factors as data type (continuous or continuous and categorical) and data category (financial or financial and non-financial) do not have an impact on accuracy prediction.*

Keywords: Deep Learning, ANN, RNN

I. INTRODUCTION

The financial sector is and always was a strong pillar of social well-being and every economy is highly dependenton it. The private sector development is, as well, built on the premises of the financial sector. It can also have an important role in providing individuals and households with monetary means for access to basic needs, such as health and education, consequently impacting poverty reduction.

Considering these elements, undeniably, there has been a great amount of research from researchers in different areas to facilitate the quality of information available in the financial sector, making financial products available,helping predict financial trends, goal evaluation, asset portfolio management, pricing IPO's, finding optimal capital structure, detecting regularities in security price movements, alleviating crediting risk by predicting default and bankruptcy, etc. In this regard, many techniques have been developed. This paper focuses on the advancements and literature background on the methods applied in bankruptcy prediction for studies.

In general, these techniques/methods can be classified in two main categories: parametric (multiple discriminant analysis (MDA), linear discriminant analysis (LDA), canonical discriminant analysis (CDA), logistic regression (LR) and Naïve Bayes (NB)) and non-parametric (artificial neural networks (ANN), support vector machine (SVM), decision trees

(DT), k-nearest neighbor (KNN), hazard models, fuzzy models, genetic algorithms (GA) and hybrid models, where multiple models are combined). Starting with the parametric models, logistic regression and discriminant analysis are some of the most used statistical techniques in empirical studies of economic phenomena. The difference between them comes from the fact that LR requires a logistic distribution. DA is mostly used for categorization or classification tasks where logistic regression is mostly used for obtaining the odds ratios for each categorization variable.

II. LITERATURE REVIEW

Using Genetic Algorithm to Improve Classification of Imbalanced Datasets for Bankruptcy Ibtissam Benchaji Samira Douzi ; Bouabid ElOuahidi IEEE 2023.

With the growing usage of Bankruptcy transactions, financial Bankruptcy crimes have also been drastically increased leading to the loss of huge amounts in the finance industry. Having an efficient Bankruptcy method has become a necessity for all banks in order to minimize such losses. In fact, Bankruptcy system involves a major challenge: the

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

Bankruptcy data sets are highly imbalanced since the number of Bankruptcy ulent transactions is much smaller than the legitimate ones. Thus, many of traditional classifiers often fail to detect minority class objects for these skewed data sets. This paper aims first: to enhance classified performance of the minority of Bankruptcy instances in the imbalanced data set, for that we propose a sampling method based on the K-means clustering and the genetic algorithm. We used K-means algorithm to cluster and group the minority kind of sample, and in each cluster we use the genetic algorithm to gain the new samples and construct an accurate Bankruptcy classifier.

An Ensemble Learning Framework for Bankruptcy Based on Training Set Partitioning and Clustering Hongyu Wang ; Ping Zhu ; Xueqiang Zou ; Sujuan Qin IEEE 2023.

The popularity of Bankruptcy has greatly facilitated the transactions between merchants and cardholders. However, Bankruptcy has been derived, which results in losses of billions of euros every year. In recent years, machine learning and data mining technology have been widely used in Bankruptcy and achieved favorable performances. Most of these studies use the technology of under-sampling to deal with the high imbalance of Bankruptcy data. However, it will potentially discard some relevant training samples which will weaken the ability of the classifier. In this paper, we propose an ensemble learning framework based on training set partitioning and clustering. It turns out that the proposed framework not only ensures the integrity of the sample features, but also solves the high imbalance of the dataset. A main feature of our framework is that every base estimator can be trained in parallel. This improves the efficiency of the framework. We show the effectiveness of our proposed ensemble framework by experimental results on a real Bankruptcy transaction dataset.

Bankruptcy on the Skewed Data Using Various Classification and Ensemble Techniques Ankit Mishra ;Chaitanya Ghorpade IEEE 2023.

Nowadays, as internet speed has increased and the prices of mobile have decreased very much in past few years. Also the data prices too are very much affordable to most of the people. This has resulted into the digitization of most of the institutes as it is easy and convenient for the people and also for the authority to maintain the records. So, it resulted in most of the banks and other institutes receiving and transferring money through credit cards. But with the hackers and other cyber criminals around there is always chances of the Bankruptcy s in the transactions. The possibility of the Bankruptcy transaction is very less but it is not negligible and even having one Bankruptcytransaction is unacceptable because it is crime and we can't neglect it even if it is very less as it harms both the customer and credibility of the institute. So this paper aims at analyzing various classification techniques using various metrics for judging various classifiers. This model aims at improving Bankruptcy rather than misclassifying a genuine transaction as Bankruptcy

Supervised Machine Learning Algorithms for Bankruptcy Bankruptcy ulent Transaction Detection: A Comparative Study Sahil Dhankhad ; Emad Mohammed ; Behrouz Far IEEE 2023.

The goal of data analytics is to delineate hidden patterns and use them to support informed decisions in a variety of situations. Bankruptcy is escalating significantly with the advancement of the modernized technology and become an easy target for Bankruptcy ulent. Bankruptcy is a severe problem in the financial service and costs billions of a dollar every year. The design of Bankruptcy algorithm is a challenging task with the lack of real- world transaction dataset because of confidentiality and the highly imbalanced publicly available datasets. In this paper, we apply different supervised machine learning algorithms to detect Bankruptcy ulent transaction using areal-world dataset. Furthermore, we employ these algorithms to implement a super classifier using ensemble learning methods. We identify the most important variables that may lead to higher accuracy in Bankruptcy ulent transaction detection. Additionally, we compare and discuss the performance of various supervised machine learning algorithms exist in literature against the super classifier that we implemented in this paper.

Bankruptcy Using Capsule Network Shuo Wang ; Guanjun Liu ; Zhenchuan Li ; Shiyang Xuan ; Chungang Yan ; Changjun Jiang IEEE 2023.

Bankruptcy is now popular in daily life. Meanwhile, Bankruptcy events occur more frequently, which result in massive financial losses. There are a number of Bankruptcy methods, but they do not deeply mine features of customer's transaction behavior so that their detection effectiveness is not too desirable. This paper focuses on two aspects of feature mining. Firstly, the features of Bankruptcy transactions are expanded in time dimension to characterize the distinct payment habits of legal users and criminals. Secondly, Capsule Network (CapsNet) is adopted to further dig some deep features on the base of the expanded features, and then a Bankruptcy model is trained to identify if a

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

transaction is legal or Bankruptcy . Through experiments on a real transaction dataset, we demonstrate that the time dimension extension can improve the performance of Bankruptcy, and then CapsNet is further illustrated to be more advantageous in Bankruptcy compared with other models.

Random forest for Bankruptcy Shiyang Xuan ; Guanjun Liu ; Zhenchuan Li ; Lutao Zheng ; Shuo Wang ;Changjun Jiang IEEE 2023.

Bankruptcy events take place frequently and then result in huge financial losses. Criminals can use some technologies such as Trojan or Phishing to steal the information of other people's credit cards. Therefore, an effective Bankruptcy method is important since it can identify a Bankruptcy in time when a criminal uses a stolen card to consume. One method is to make full use of the historical transaction data including normal transactions and Bankruptcy ones to obtain normal/Bankruptcy behavior features based on machine learning techniques, and then utilize these features to check if a transaction is Bankruptcy or not. In this paper, two kinds of random forests are used to train the behavior features of normal and abnormal transactions. We make a comparison of the two random forests which are different in their base classifiers, and analyze their performance on credit Bankruptcy .The data used in our experiments come from an e-commerce company in China.

A utilitarian approach to adversarial learning in Bankruptcy Tyler Cody ; Stephen Adams ; Peter A. Beling IEEE 2023.

Bankruptcy can be modeled as an adversarial game between a Bankruptcy ster and the Bankruptcy mechanism. Previous work uses a game theoretical adversarial learning approach to model the most successful strategy and preemptively adapt the Bankruptcy system. The game consists of the adversary's selection of strategy and the Bankruptcy system's decision to retrain. In the previous work, the detection system is retrained every round of the game. In application, there may be costs and risks associated with training and deploying a new model. Thus,it may be desirable to optimize the decision of whether to retrain the model based on the expected economic disutility. The presented work addresses this desire by using a utilitarian approach to optimally decide whether to retrain the classifier by comparing the economic values of the new and old classifiers. A framework from decisiontheory is derived within the context of Bankruptcy. Further, we show how a utility function can be used to identifythe best Bankruptcy strategy in economic terms. We add to the literature by extending the adversarial learning model developed in previous work to include a theoretical framework for retraining when it is economically advantageous and judging Bankruptcy strategies on their economic cost. Our approaches are tested against the decisions to always retrain and to never retrain.

Bankruptcy using autoencoder based clustering Mohamad Zamini ; Gholamali Montazer IEEE 2023.

By increasing growth of e-commerce, which has coupled with the increase in online payments, Bankruptcy has become an important issue for banks. Bankruptcy in financial transactions can cause heavy damages and endanger their reputation among their customers. Thus, focusing on a variety of Bankruptcy methods, as well as new ways to tackle and preventing them, is becoming increasingly important. In this paper, we have proposed an unsupervised Bankruptcy method using autoencoder based clustering. An autoencoder with three hidden layer and a k-means clustering has been used and tested on 284807 transactions from European banks. Based on the results, the accuracy of this method was 98.9%, as well as 81% TPR which outperforms in comparison with others..

Bankruptcy : A Realistic Modeling and a Novel Learning Strategy Andrea Dal Pozzolo ; Giacomo Boracchi ;Olivier Caelen ; Cesare Alippi ; Gianluca Bontempi IEEE 2023.

Detecting Bankruptcy s in Bankruptcy transactions is perhaps one of the best testbeds for computational intelligence algorithms. In fact, this problem involves a number of relevant challenges, namely: concept drift (customers' habits evolve and Bankruptcy sters change their strategies over time), class imbalance (genuine transactions far outnumber Bankruptcy s), and verification latency (only a small set of transactions are timely checked by investigators). However, the vast majority of learning algorithms that have been proposed for Bankruptcy rely on assumptions that hardly hold in a real-world Bankruptcy -detection system (FDS). This lack of realism concerns two main aspects: 1) the way and timing with which supervised information is provided and2) the measures used to assess Bankruptcy -detection performance. This paper has three major contributions. First, we propose, with the help of our industrial partner, a formalization of the Bankruptcy -detection problem that realistically describes the operating conditions of FDSs that everyday analyze massive streams of Bankruptcy transactions. We also illustrate the most appropriate performance measures to be used for Bankruptcy -detection purposes. Second, we design and assess a novel learning strategy that

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

effectively addresses class imbalance, concept drift, and verification latency. Third, in our experiments, we demonstrate the impact of class unbalance and concept drift in a real-world data stream containing more than 75 million transactions, authorized over a timewindow of three years.

III. METHODOLOGY SECTION

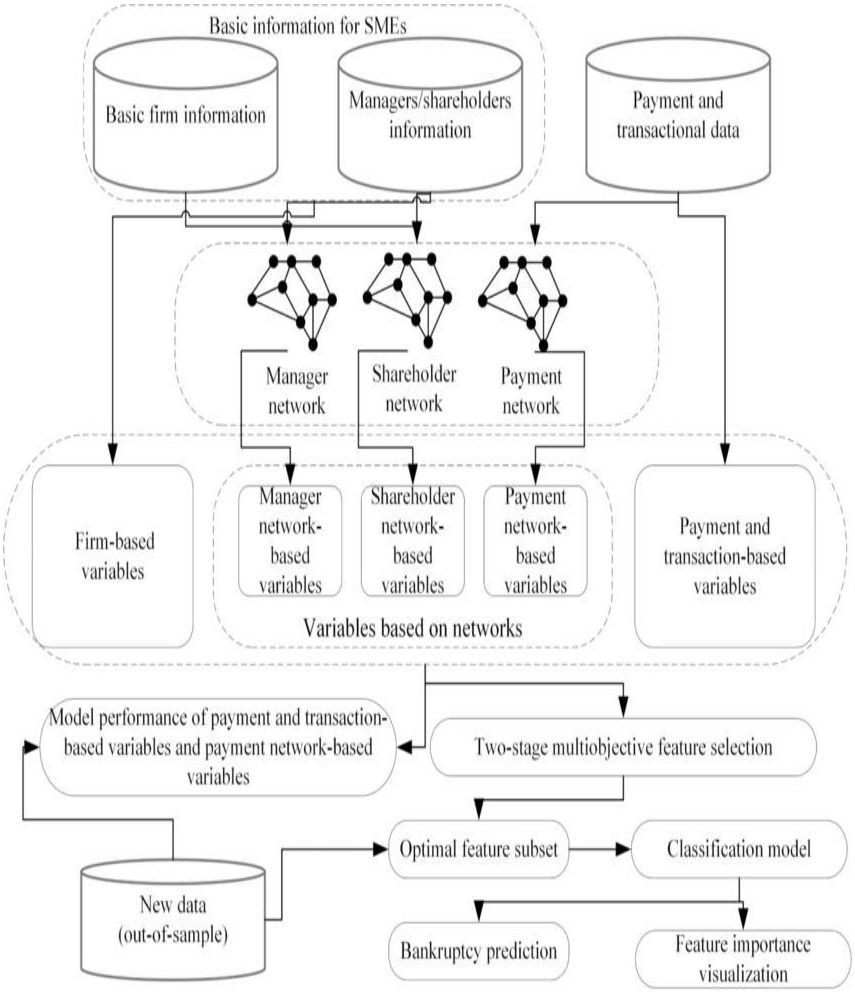

Fig 1: System Architecture

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

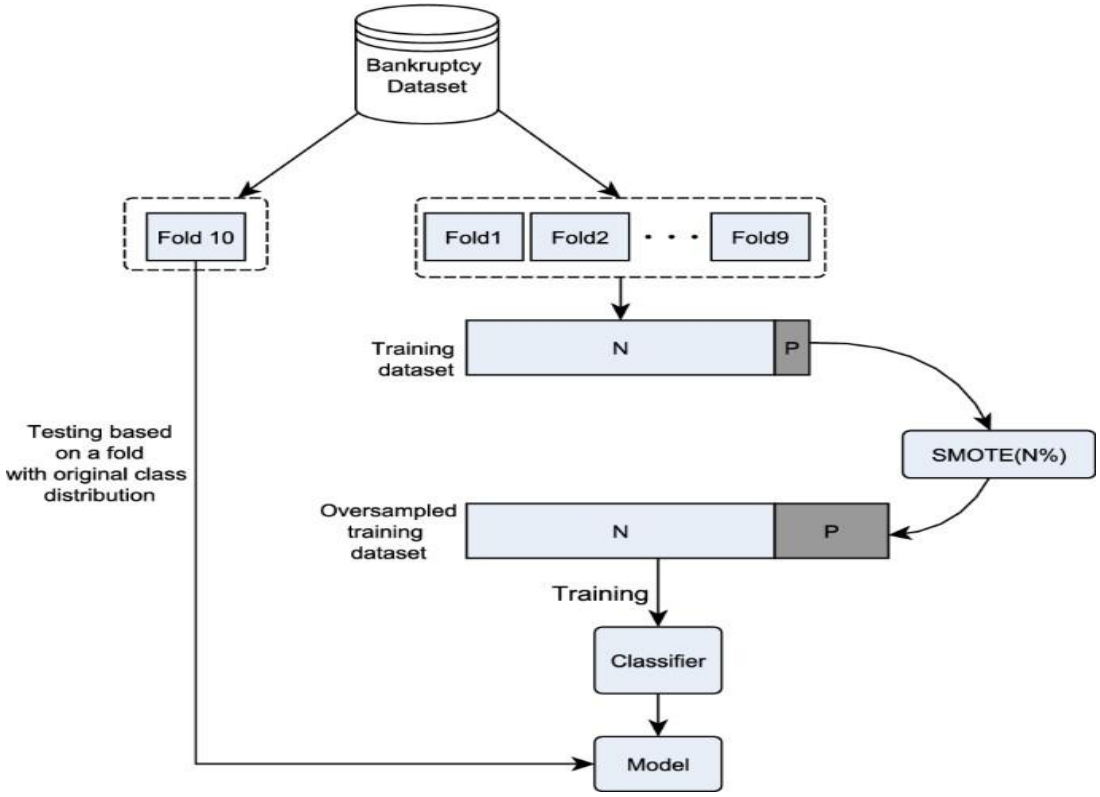

Fig 2 : UML Diagram

IV. EXPERIMENTAL RESULTS

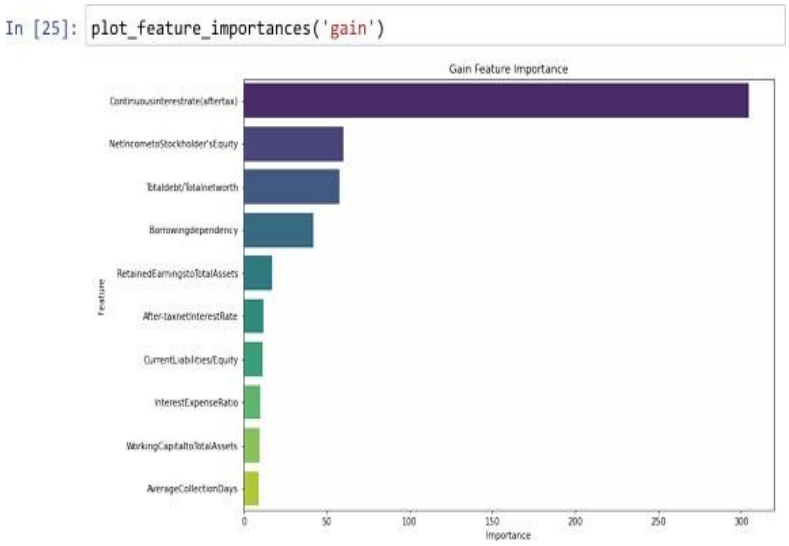

Fig3 Gain feature importance

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

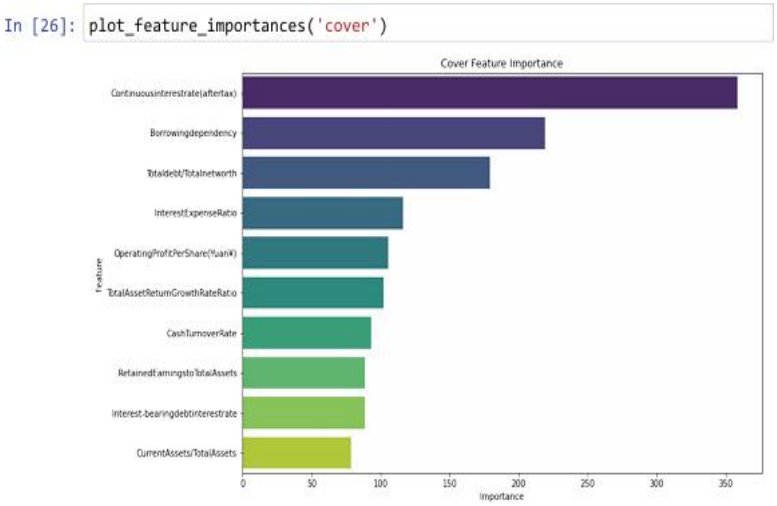

Fig.4 Feature importance Cover

V. CONCLUSION

Application permissions are significant in Android operating system security. These permissions, which are extracted from applications, are used as attributes to detect malicious software with machine learning algorithms in this study. Android malware detection is carried out with two rule-based classification models using Hybrid models. However, classifiers are quite simple and easy to use. This is the most significant advantage of the proposed approaches.

VI. FUTURE ENHANCEMENT

Future research should consider other machine learning algorithms to ascertain more efficient ways to perform the classification technique on the datasets. It is recommended that further research should be carry out on other parameters that can improve the accuracy of detection.

REFERENCES

[1]. F. Akyildiz et al., "Wireless Sensor Networks: A Survey, "Elsevier Comp. Networks, vol. 3, no. 2, 2019, pp. 393–422

[2]. G.Li, J.He, Y. Fu. "Group-based Malware detection system in wireless sensor networks" Computer Communications, Volume 31, Issue 18 (December 2019)

[3]. Michael Brownfield, "Wireless Sensor Network Denial of Sleep Attack", Proceedings of the 2019 IEEE Workshop on Information Assurance and Security United States Military Academy,West Point, NY.

[4]. FarooqAnjum, DhanantSubhadrabandhu, SaswatiSarkar *, Rahul Shetty, "On Optimal Placement of Malware Detection Modules in Sensor Networks", Proceedings of the First International Conference on Broadband Networks (BROADNETS19).

[5]. Parveen Sadotra et al, International Journal of Computer Science and Mobile Computing, Vol.5 Issue.9, September- 2019, pg. 23-28

[6]. K. Akkayaand M. Younis, —A Survey of Routing Protocols in Wireless Sensor Networks, in the Elsevier Ad Hoc Network Journal, Vol. 3/3 pp. 325-349, 2019.

[7]. AAbduvaliyev, S. Lee, Y.K Lee, "Energy Efficient Hybrid Malware Detection System for Wireless Sensor Networks", IEEE International Conference on Electronics and Information Engineering, Vol.2, pp. 25-29, August 2019.

[8]. Parveen Sadotra and Chandrakant Sharma. A Survey: Intelligent Malware Detection System in Computer Security. International Journal of Computer Applications 151(3):18-22, October 2019

[9]. A Araujo, J. Blesa, E. Romero, D. Villanueva, "Security in cognitive wireless sensornetworks. Challenges and open problems", EURASIP Journal on Wireless Communications andNetworking, February 2019.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 7, May 2024

[10]. A. Becher, Z. Benenson, and M. Dorsey, \Tampering with motes: Real-world physical attacks on wireless sensor networks." in SPC (J. A. Clark, R. F. Paige, F. Polack, and P. J.Brooke,eds.), vol. 3934 of Lecture Notes in Computer Science, pp. 104{118, Springer, 2019.

[11]. Krontiris and T. Dimitriou, \A practical authentication scheme for in-network programming in wireless sensor networks," in ACM Workshop on Real- World Wireless SensorNetworks, 2019.

[12]. M. Ali Aydın *, A. HalimZaim, K. GokhanCeylan "A hybrid Malware detection system design for computer network security" Computers and Electrical Engineering 35 (2019) 517–526