IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 4, April 2024

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304282 |

An Intelligent Data-Driven Model to secure Intra vehicle Communications Based on Machine Learning

Dr. M. Anantha Guptha[1], Byalla Maheswari[2], Bangi Naresh[3], S Mohammed Imran Nazeer[4], Annam Pavan Sai[5]

UG Student, Dept. of E.C.E., GATES Institute of Technology, Gooty, Anantapur, Andhra Pradesh, India[2,3,4,5]

Assistant Professor, Dept. of E.C.E., GATES Institute of Technology, Gooty, Anantapur, Andhra Pradesh, India [1]

ABSTRACT: The high relying of electric vehicles on either in vehicle or between-vehicle communications can cause big issues in the system. This paper is going to mainly address the cyber-attack in electric vehicles and propose a secured and reliable intelligent framework to avoid hackers from penetration into the vehicles. The proposed model is constructed based on an improved support vector machine model for anomaly detection based on the controller area network (CAN) bus protocol. In order to improve the capabilities of the model for fast malicious attack detection and avoidance, a new optimization algorithm based on social spider (SSO) algorithm is developed which will reinforce the training process at offline. Also, a two-stage modification method is proposed to increase the search ability of the algorithm and avoid premature convergence. Last but not least, the simulation results on the real data sets reveal the high performance, reliability and security of the proposed model against denial-of-service (DoS) hacking in the electric vehicles

I. INTRODUCTION

Technically, vehicles are composed of many hardware modules namely called electronic control units (ECUs) being controlled by different software tools. All sensors installed in a vehicle will send their data to the ECU, where this data are processed and the requiring orders are sent to the relevant actuators [1]. Such a highly complex hardware software data transfer process may happen through the use of different network protocols such as CAN, LIN, Flex Ray or MOST [2]. Among these protocols, CAN bus is the most popular one not only in vehicles, but also in medical apparatuses, agriculture, etc due to its high capability and promising characteristics.

Some of the main advantages of the CAN bus standard may be briefly named as allowing up to 1Mbps data rate transfer, reducing the wiring in the device saving cost and time due to the simple wiring, auto retransmission of lost messages and error detection capability [3]. Unfortunately, since CAN bus protocol was devised at a time where vehicles were almost isolated, this standard suffers from some security issues in the new dynamic environment of smart grids. This will motivate the hackers to attack the electric vehicles through the ECU and inject malicious messages into their systems.

In [4], some cyber intrusion scenarios are modelled and applied on the electric vehicles to assess their vulnerabilities and possible side effects getting finally into the power grid. In [5], a new classification method is developed for cyber intrusion detection in vehicles. In [6], a data intrusion detection system is developed which can detect the cyber-attack based on the CAN bus message frequency increase or CAN message ID misuse. This will help the driver to detect that an attack has happened so to stop the vehicle immediately. In [7], authors suggest that all CAN messages should pass a data management system to avoid any cyber intrusion.

In [8], an algorithmic solution is used to stop attacks of types of denial-of-service or errorflag in the vehicle. In [9], it is suggested to assign an ECU as the master ECU in the manufacturing stage of the vehicle so to run an attestation process in the system. In [10], a firewall is introduced for the vehicle to sit between the CAN bus and the communicating system and stop the cyber attack commands to the CAN bus. In [11], an intrusion detection system based on the traffic entropy of in-vehicle network communication system of the CAN, to propose an intelligent and highly secure method to equip the electric vehicles with a powerful anomaly detection and avoidance mechanism. The

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304282 |

proposed method is constructed based on support vector machine and the concept of one-class detection system to avoid any malicious behaviour in the vehicle [13].
Here the experimental CAN bus data are used to let the support vector machine learn the normal frequency of the different message frames at different commands. In order to get into the maximum capability of the model, a new optimization algorithm based on social spider optimization (SSO) algorithm is proposed to adjust the SVR setting parameters, properly [14].

Due to the high complexity and nonlinearity of the electric vehicle CAN bus dataset, a new two-stage modification method based on crossover and mutation operators of geneticalgorithm is developed which can increase the algorithm population diversity and at the same time avoid premature convergence. The feasibility and satisfying performance of the proposed model are examined using the real datasets gathered from an electric vehicle.

KEYWORDS: Anomaly detection, controller area networks (CAN) bus, intra vehicle.

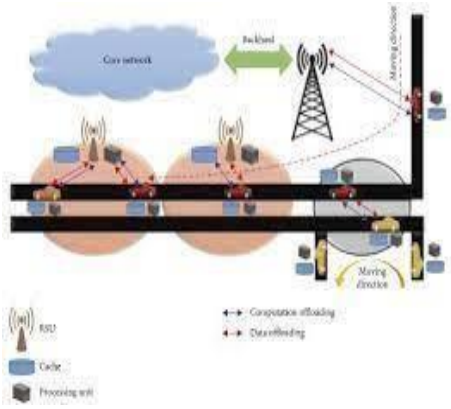
II. EXISTING SYSTEM

In electric vehicles, CAN standard is the most widely used protocol by automakers for communications with low cost in the units with a high number of components, up to 500 million chips. In the vehicle industry, the CAN resiliency and noise resistance level is acceptable owing to its structure. Unfortunately, CAN bus protocols do not offer confidentiality and authenticationto CAN data frames so making it possible for hackers to enter the vehicle system, either on a wired or wireless approach.

In the wired approach, one can communicate with the CAN bus through the OBD-II maintenance port located under the steering in most vehicles. Although the main idea behind thisport is to be used for diagnostics of engine and vehicle maintenance, but it will let hackers take the CAN packets using a simple scanning tool.

From this point, it is easy to read and write traffic in the CAN bus with the use of ECOM API such as CAN Receive Message and CAN Transmit Message [10]. In the wireless attack, the cyber interfere is the same by targeting ECU except that the penetration point is not OBD-II. While the penetration points in the wireless hacking can be different, but in most of them it is required for the car to be connected to a malicious WIFI hotspot. Also, the security mechanismof the transponder can be reverse engineered in the keyless vehicles.

The research reveals several weaknesses in the design of the cipher, the authentication protocol and also in their implementation. Some of the other wireless entry points to vehicles can be named as ―Wireless connection between sensors and ECUs such as TPMS system Add-on technologies, entertainment system (gaming), smart key‖ and ―Internet, smart infrastructures‖.



System architecture

DISADVANTAGES OF EXISTING SYSTEM:
Less accuracy
low Efficiency

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304282 |

III. PROPOSED SYSTEM

The last sections were mainly focusing on the proposed model, the theories and backgrounds. In this section, the performance of the proposed model is examined using the experimental data gathered from an electric car. This paper assesses the DoS attack since it is focusing on the vehicle intra-communication within which DoS has a high significance among different attacks. In the DoS attack, the hacker attempts to prevent legitimate users (driver) from accessing the service. Considering the fact that vehicles are mobile devices, DoS attack is so dangerous (and thus important) in vehicles since it can make severe car crash or losses.

Examples of hackings achieved through the DoS attack in the vehicles are activating the brakes while the vehicle is in motion, turning the steering wheel to the left/right suddenly, turning off the engine, unlocking a door, etc. According to the analysis from the recorded CAN traffic during a normal driving time of 10-minute, each message frame with a specific ID has some unique frequencies which can be learned by the proposed anomaly detection model. In Table II, a set of CAN bus identifiers.

In order to make sure that our model is learning all possible ID numbers, we had a complete trace analysis.
Therefore, after capturing the traffic log and implementing the trace analysis, it was realized that a 10-min driving scenario would capture the majority of the messages that are occurring commonly, owing to the fact that most of the CAN messages are periodic. Therefore, the model developed can be regarded as the proof-of-concept that shows the proposed anomaly detection model can learn the existing pattern in the CAN messages to distinguish between normal and anomalous behaviors in the testing phase.

In order to make a realistic condition for the driving test, the CAN traffic file covers the following conditions: the engine ignition was turned on and the vehicle remained at a standstill for a few seconds and then the gear was engaged to ―D‖ mode. Then, the vehicle is driven for about 8 minutes at a public street. For several times, the brake pedal is also pressed during the drive.

The car is then stopped and the gear mode is changed to ―R‖ to drive backwards a bit and make a parking manoeuvre. Finally, the gear moves to ―P‖ mode so the vehicle would remain at the standstill for a few seconds and then the engine is turned off.

ADVANTAGES OF PROPOSED SYSTEM:
High accuracy
High efficiency

In this manner, subsequent to catching the traffic log and executing the follow examination, it was understood that a 10-min driving situation would catch most of the messages that are happening usually, attributable to the way that the vast majority of the CAN messages are occasional. Subsequently, the model created can be viewed as the verification of-idea that shows the proposed abnormality location model can gain proficiency with the current example in the CAN messages to recognize typical and strange ways of behaving in the testing stage.

The car is then stopped and the gear mode is changed to "R" to drive backwards a bit and make a parking manoeuvre. Finally, the gear moves to "P" mode so the vehicle would remain at the standstill for a few seconds and then the engine is turned off The feasibility of the project is analysed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company.

IV. LITERATURE SURVEY

SYSTEM STUDY FEASIBILITY STUDY
The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

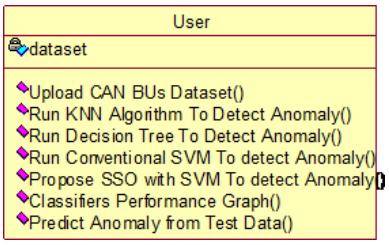
Three key considerations involved in the feasibility analysis are

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304282 |

ECONOMICAL FEASIBILITY
TECHNICAL FEASIBILITY
SOCIAL FEASIBILITY
ECONOMICAL FEASIBILITY:



This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

TECHNICAL FEASIBILITY:
This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system musthave a modest requirement, as only minimal or null changes are required for implementing this system.

SOCIAL FEASIBILITY:
The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

SYSTEM DESIGN
UML stands for Unified Modelling Language. UML is a standardized general-purpose modelling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.
The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta- model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modelling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modelling and other non-software systems.
The UML represents a collection of best engineering practices that have proven successful in themodeling of large and complex systems.

GOALS:
The Primary goals in the design of the UML are as follows:
Provide users a ready-to-use, expressive visual modelling Language so that they can developand exchange meaningful models.
Provide extendibility and specialization mechanisms to extend the core concepts.
Be independent of particular programming languages and development process.
Provide a formal basis for understanding the modelling language.
Encourage the growth of OO tools market.
Support higher level development concepts such as collaborations, frameworks, patterns andcomponents.
Integrate best practices.

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

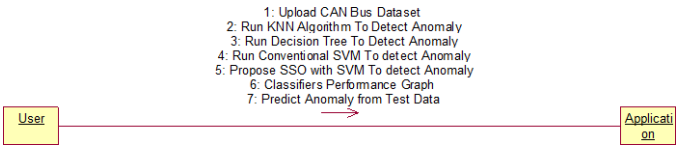
| DOI:10.15680/IJIRSET.2024.1304282 |

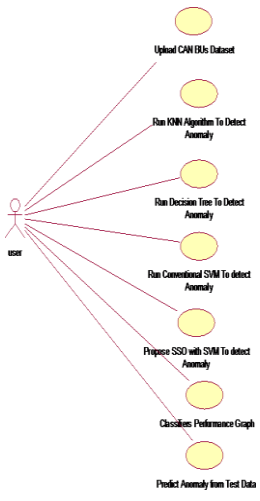
V. NETWORK SETUP

USE CASE DIAGRAM:
A use case diagram in the Unified Modelling Language (UML) is a type of behavioural diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

CLASS DIAGRAM:
In software engineering, a class diagram in the Unified Modelling Language (UML) is atype of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.

COLLABRATION DIAGRAM:
Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modelling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system





| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304282 |

COLLABRATION DIAGRAM:
Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

VI. IMPLEMENTATION

MODULES:
Upload CAN Bus Dataset ‘button and upload dataset
Run KNN Algorithm to Detect Anomaly ‘button to build KNNclassifier train model to detect anomaly and evaluate its performancebased on 4 indices
Run Conventional SVM To detect Anomaly ‘button to evaluateconventional SVM performance.
Propose SSO with SVM To detect Anomaly ‘button to run propose SSO with SVM classifier and evaluate its performance.
Classifiers Performance Graph ‘button to get performance graph between allclassifiers
Predict Anomaly from Test Data ‘button to upload test data and predict itlabel

SOFTWARE ENVIRONMENT:
4. 1. Tensor flow:
TensorFlow is a free and open-source software library for dataflow and differentiable programming_across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. It is used for both research and production at Google. TensorFlow was developed by the Google Brain team for internal Google use. It was released under the Apache 2.0 open-source license on November 9, 2015.

4. 2. NumPy:
NumPy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays. It is the fundamental package for scientific computing with Python. It contains various features including these important ones:

- A powerful N-dimensional array object
- Sophisticated (broadcasting) functions
- Tools for integrating C/C++ and Fortran code
- Useful linear algebra, Fourier transform, and random number capabilities

Besides its obvious scientific uses, NumPy can also be used as an efficient multi-dimensional container of generic data. Arbitrary data-types can be defined using NumPy which allows NumPy to seamlessly and speedily integrate with a wide variety of databases.

4. 3 Pandas:
Pandas is an open-source Python Library providing high-performance data manipulation and analysis tool using its powerful data structures. Python was majorly used for data munging and preparation. It had very little contribution towards data analysis. Pandas solved this problem. Using Pandas, we can accomplish five typical steps in the processing and analysis of data, regardless of the origin of data load, prepare, manipulate, model, and analyse. Python with Pandas is used in a wide range of fields including academic and commercial domains including finance, economics, Statistics, analytics, etc.

VII. HARDWARE AND SOFTWARE REQUIREMENTS

HARDWARE REQUIREMENTS
System: i3 or above
RAM: 4GB
Hard Disk: 40GB

SOFTWARE REQUIRMENTS
Operating system: Windows8 or Above
Coding language: Python

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304282 |

VIII. RESULTS AND ANALYSIS

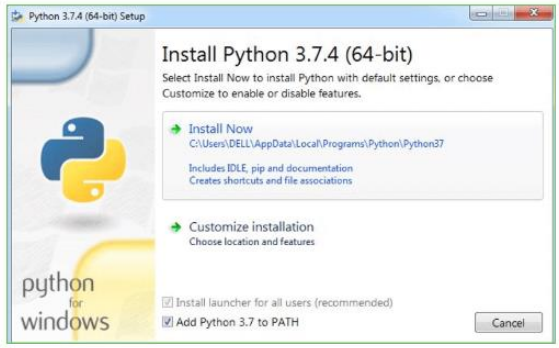


FIG: 9.1 Click on Install NOW & Click on Close

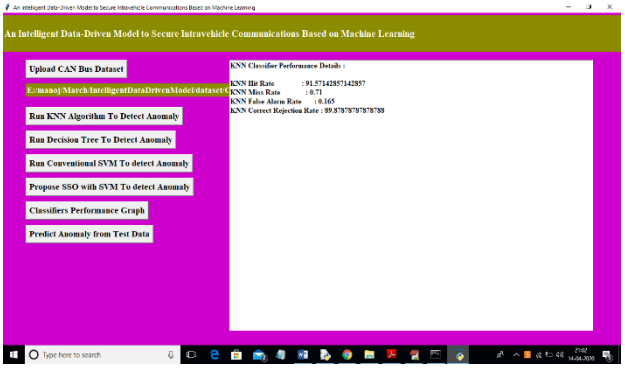


FIG: 9.2 Upload Dataset

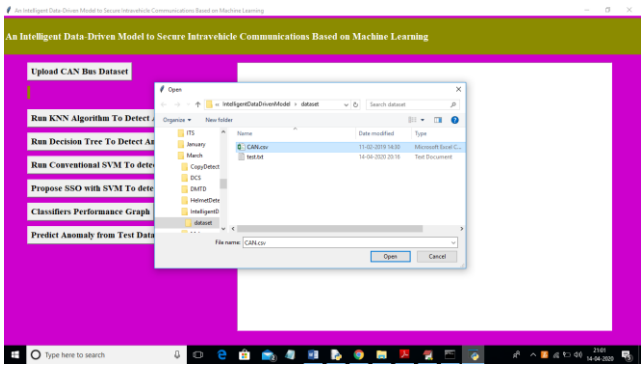


FIG: 9.3 Select Dataset

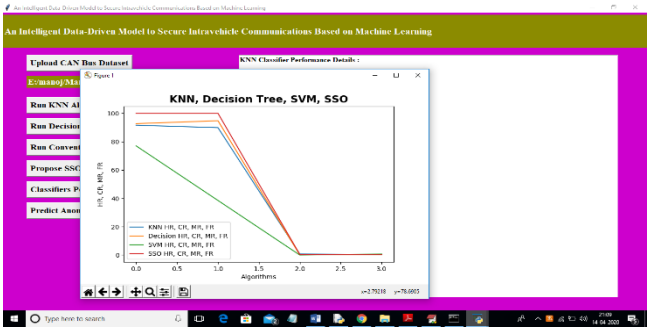


FIG: 9.4 Select KNN Algorithm

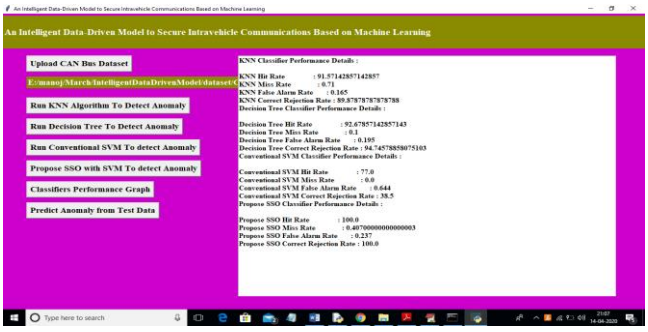


FIG: 9.5 SVM Algorithm

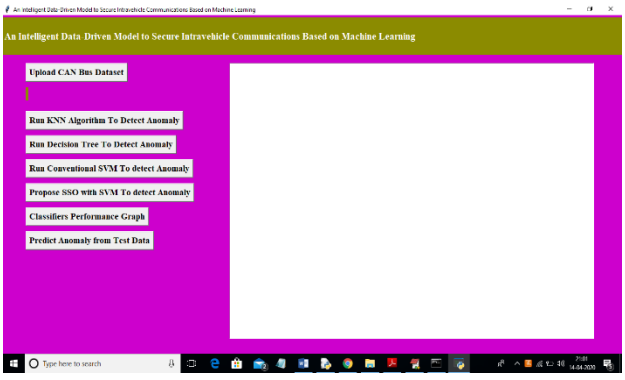


FIG: 9.6 Comparison Graph

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304282 |

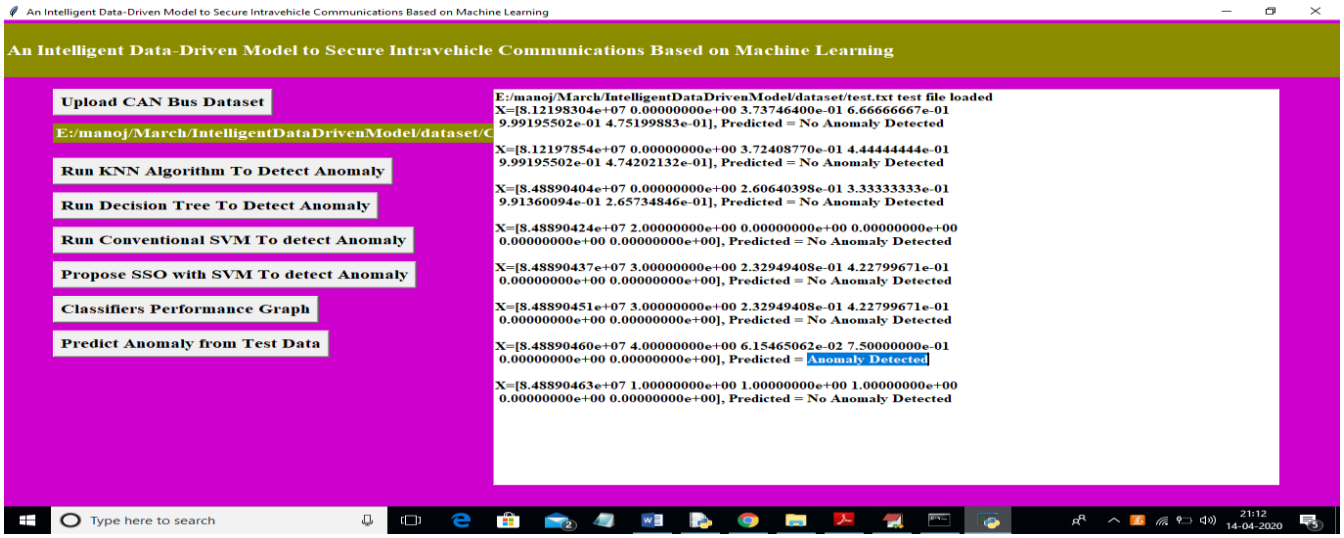


FIG: 9.7 All Data Sets

To download Windows 32-bit python, you can select any one from the three options: Windows x86 embeddable zip file, Windows x86 executable installer or Windows x86 web-based installer. To run project double, click on 'run.bat' file to get below screen. In above screen click on 'Upload CAN Bus Dataset' button and upload dataset. In above screen I am uploading 'CAN.csv' dataset and after uploading dataset. Now click on 'Run KNN Algorithm o Detect Anomaly' button to build KNN classifier train model to detect anomaly and evaluate its performance based on 4 indices. In above screen we got 4 indices values for KNN algorithm and now click on 'Run Decision Tree to Detect Anomaly' button to evaluate decision tree performance. In above screen we got SVM performance data and now click on 'Propose SSO with SVM to detect Anomaly' button to run propose SSO with SVM classifier and evaluate its performance. In above screen for SSO we got performance metric as 100% and MR and FR is not mandatory so we can ignore as said in paper. Now click on 'Classifiers Performance Graph' button to get performance graph between all classifiers. In above graph propose SSO has given high performance compare to other algorithms. In above graph y-axis represents HR, MR, FR and CR values. Now click on 'Predict Anomaly from Test Data' button to upload test data and predict it label. In above screen in text area we can see uploaded test data and its predicted class label. All records contain normal packet data accept one record. So, by using machine learning algorithms we can analyse packets and if packet contains attack, then we ignore processing such packets.

IX. CONCLUSION

This paper proposed a novel intelligent and secured anomaly detection model for cyber attack detection and avoidance in the electric vehicles. The proposed model is constructed based on an improved support vector machine model reinforced by the MSSO algorithm. From the cyber security point of view, the proposed model could successfully detect malicious behaviors while letting the trusted message frames broadcast in the CAN protocol. The high HR% and FR% indices prove the true positive and true negative decisions made by the proposed model. Regarding the MR% and CR% indices, the very low values which most of them are around the upper and lower bounds of the message frame frequency, show the highly trustable performance of this model. The authors will assess the effect of other cyberattacks on the performance of different anomaly detection models in the future works.

REFERENCES

1.A. Monot; N. Navet; B. Bijoux; F. Simo not-Lion, ―Multisource Software on Multicore Automotive ECUs—Combining Runnable Sequencing With Task Scheduling‖, IEEE Trans. Industrial Electronics, vol. 59, no. 10. Pp. 3934-3942, 2012.
2.T.Y. Moon; S.H. Seo; J.H. Kim; S.H. Hwang; J. Wook Jeon, ―Gateway system with diagnostic function for LIN, CAN and Flex Ray‖, 2007 International Conference on Control,Automation and Systems, pp. 2844 – 2849, 2007.

| e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 13, Issue 4, April 2024 ||

| DOI:10.15680/IJIRSET.2024.1304282 |

3.B. Groza; S. Murvay, –Efficient Protocols for Secure Broadcast in Controller Area Networks‖, IEEE Trans. Industrial Informatics, vol. 9, no. 4, pp. 2034-2042, 2013.
4.B. Mohandes, R. Al Hammadi, W. Sanusi, T. Mezher, S. El Khatib, –Advancing cyber– physical sustainability through integrated analysis of smart power systems: A case study on
electric vehicles‖, International Journal of Critical Infrastructure Protection, vol. 23, pp. 33-48,2018.
5.G. Loukas, E. Karapistoli, E. Panaousis, P. Sarigiannidis, T. Vuong, A taxonomy and survey of cyber-physical intrusion detection approaches for vehicles, Ad Hoc Networks, vol. 84, pp. 124-147, 2019.
6.Hoppe T, Kiltz S, Dittmann J. Security threats to automotive can networks. practical examples and selected short-term countermeasures. Reliab Eng Syst Saf vol. 96, no. 1, pp. 11–25, 2011.
7.Schulze S, Pukall M, Saake G, Hoppe T, Dittmann J. On the need of data management in automotive systems. In: BTW, vol. 144; pp. 217–26, 2009.
8.Ling C, Feng D. An algorithm for detection of malicious messages on can buses. 2012national conference on information technology and computer science. Atlantis Press; 2012.
9.Oguma H, Yoshioka X, Nishikawa M, Shigetomi R, Otsuka A, Imai H. New attestation-based security architecture for in-vehicle communication. In: Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE. IEEE; pp. 1–6, 2008.
10.L. Pan, X. Zheng, H. X. Chen, T. Luan, L. Batten, ―Cyber security attacks to modern vehicular systems‖, Journal of Information Security and Applications, vol. 36, pp. 90-100, October 2017.
11.Kang, M. J., & Kang, J. W., –Intrusion detection system using deep neural network for in- vehicle network security‖, PloS one, vol. 11, no. 6, e0155781, 2016.
12.Theissler, A., –Detecting known and unknown faults in automotive systems using ensemble-based anomaly detection‖, Knowledge-Based Systems, vol. 123, pp. 163-173.
13.F. Zhu, J. Yang, C. Gao, S. Xu, T. Yin, –A weighted oneclass support vector machine‖, Neurocomputing, vol. 189, pp. 1-10, 12 May 2016.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details