International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 09 | Sep 2023 www.irjet.net p-ISSN: 2395-0072

Prediction of Wireless Sensor Network and Attack using Machine Learning Technique

M. Nirmala Devi[1], V. Pavithra[2]

[1]*Student, M.E, Dept. of Computer Science and Engineering, T.J.S Engineering College, Tamil Nadu, India*
[2]*Assistant Professor, Dept. of Computer Science and Engineering,T.J.S Engineering College, Tamil Nadu, India*

Abstract - *Wireless sensor network has attracted significant attention in research and development due to its tremendous applications in medical, military and defence, medical, environmental, industrial, infrastructure protection, and commercial applications to enable to interact with each other controlled remotely. A Wireless Sensor Network (WSN) has wide applications such as environmental monitoring and tracking of the target nodes for communication. The sensor nodes are equipped with wireless interfaces used for communication between the nodes and another network. Wireless Sensor Network suffers from many constraints that make security a primary challenge. When the sensor node is deployed in a communication environment unattended, the nodes are vulnerable to various attacks. The analysis of dataset by supervised machine learning technique(SMLT) to capture several information's like, variable identification, univariate analysis, bivariate and multivariate analysis, missing value treatments etc. A comparative study between machine learning algorithms had been carried out in order to determine which algorithm is the most accurate in predicting the type WSN attacks. The results show that the effectiveness of the proposed machine learning algorithm technique can be compared with best accuracy, precision, Recall, F1 Score, Sensitivity, and Specificity.*

Key Words: Anaconda Navigator, Jupyterlab, Spyder, PyCharm, Orange 3 app, Jupyter Notebook, RStudio.

1.INTRODUCTION

Wireless Sensor Network (WSN) is one of the best strategies for some continuous applications, because of its minimization, cost-viability, and simplicity of sending. The capability of the WSN is to screen the field of interest, gather the information, and send it to the base station (Passage) for post-handling examination. Countless sensor hubs are utilized in some WSN executions. What's more, these remote hubs have a restricted battery duration and memory limit. In this way, to get the most out of these WSNs, there should be an administration framework for these WSN hubs fit for directing the relationship among themselves and with the passageway too.

For instance, the ZigBee and 6LoWPAN are two conventions that help the board in WSNs created by the Web Designing Group (IETF) for standard transmission over IEEE 802.15.4. These conventions support current administration frameworks to utilize IEEE 802.15.4 in the 2.4 GHz band and backing short transmission. For instance, 6LoWPAN IPv6 gives an association between WSNs in view of IP tends to be on various layers. It additionally utilizes the 6LoWPAN Low Power and Misfortune Organization (RPL) standard to plan the organization geography and utilizations the AES encryption calculation to get the WSN association. Be that as it may, as the geography of these kinds of organizations is continually transforming, it will affect network directing systems, delay, multi-facet plan, inclusion, Nature of Administrations (QoS), and shortcoming recognition. In this way, it is important to rethink the administration of WSNs by planning or consolidating new conventions to manage the idea of the conditions for which these implanted gadgets are planned.

Many overviews talked about the job of AI calculations in different fields of wireless sensor networks and the Internet of Things (IoT). Moreover, ML calculations enjoy an extraordinary benefit in breaking down bundles as they travel between WSN hubs and recognizing dubious hubs.

2. EXISTING SYSTE

In a Wireless Sensor Networks (WSN) based fluid pipeline leak monitoring system, numerous sensors are deployed along the pipeline networks. A great number of measurements are continuously transmitted from the sensor nodes to their corresponding sink nodes. The energy consumed on data transmission dominates the pressure wave propagation speed, can be online updated, resulting in improvement of the leak localization accuracy. power depletion of a WSN system. To reduce the amount of data transmission and prolong the lifetime of WSN, in this paper, a Combined Dual-Prediction based Data Fusion (CDPDF) method is proposed. Transmissions are only triggered if the measurement is substantially different from the predicted value. Furthermore, unlike existing methods which establish the predictor by merely considering the measurements from a single sensor, the proposed CDPDF learns and updates the predictor by integrating measurements from multiple neighboring sensors, hence the spatial cross-correlation is taken into account and the prediction accuracy is significantly improved. In this paper, an Enhanced Leak Detection and Isolation (EnLDI) method is also proposed in

which several important parameters, such as the friction factor and the

3. PROPOSED WORK

The proposed model is to build a machine learning model for predicting wsn attacks. Previously they finds the accurate leak detection and isolation results only. wsn attack detection is an important technique for recognizing fraud activities, suspicious activities, network intrusion, and other abnormal events that may have great significance but are difficult to detect. The machine learning model is built by applying proper

data science techniques like variable identification which is the dependent and independent variables. Each and every column's features are analyzed. Then the pre-processing and visualization of the data are done. The model is built based on the previous dataset where the algorithm learns data and gets trained different algorithms are used for better comparisons. The performance metrics are calculated and compared.

4. METHODOLOGY

4. 1 Preparing Dataset

This Dataset contains 500 records. It is classified into 5 classes.

- Blackhole
- Flooding
- Grayhole
- Normal
- Scheduling

5. ANACONDA NAVIGATOR

Anaconda Navigator is a desktop graphical user interface (GUI) included in Anaconda® distribution that allows you to launch applications and easily manage conda packages, environments, and channels without using command-line commands. Navigator can search for packages on Anaconda.org or in a local Anaconda Repository.

Anaconda. Now, if you are primarily doing data science work, Anaconda is also a great option. Anaconda is created by Continuum Analytics, and it is a Python distribution that comes preinstalled with lots of useful python libraries for data science.

Anaconda is a distribution of the Python and R programming languages for scientific computing (data science, machine learning applications, large-scale data processing, predictive analytics, etc.), that aims to simplify package management and deployment.

In order to run, many scientific packages depend on specific versions of other packages. Data scientists often use multiple versions of many packages and use multiple environments to separate these different versions.

The command-line program conda is both a package manager and an environment manager. This helps data scientists ensure that each version of each package has all the dependencies it requires and works correctly.

Navigator is an easy, point-and-click way to work with packages and environments without needing to type conda commands in a terminal window. You can use it to find the packages you want, install them in an environment, run the packages, and update them – all inside Navigator.

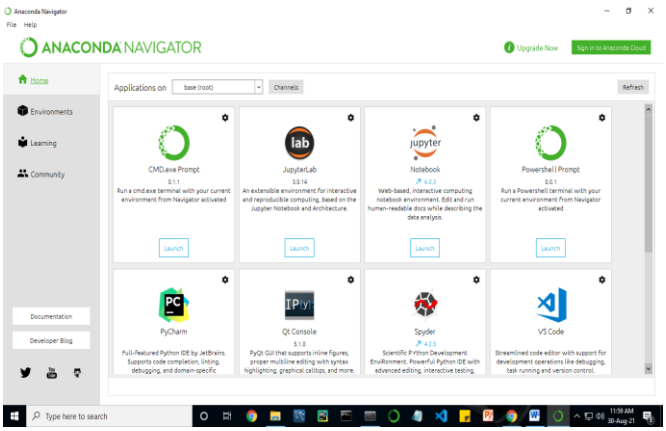

Fig-5:Anaconda Navigator Application 1

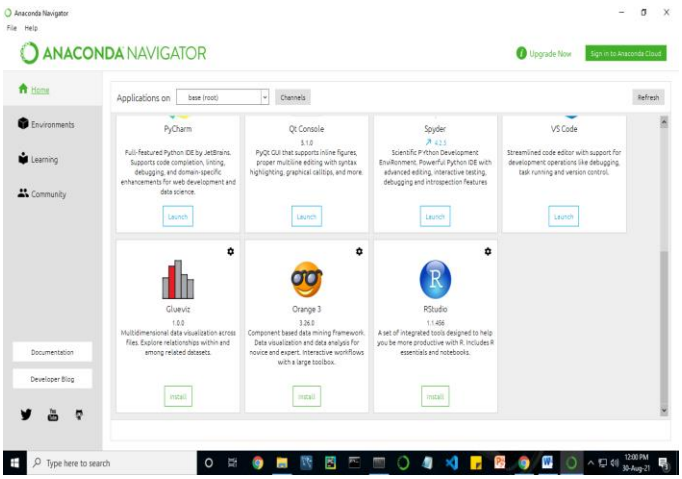

FIG-5.1:Anaconda Navigator Application 2

6. JUPYTER NOTEBOOK

This website acts as "meta" documentation for the Jupyter ecosystem. It has a collection of resources to

navigate the tools and communities in this ecosystem, and to help you get started.

Project Jupyter is a project and community whose goal is to "develop open-source software, open-standards, and services for interactive computing across dozens of programming languages". It was spun off from IPython in 2014 by Fernando Perez.

Notebook documents are documents produced by the Jupyter Notebook App, which contain both computer code (e.g. python) and rich text elements (paragraph, equations, figures, links, etc...). Notebook documents are both human-readable documents containing the analysis description and the results (figures, tables, etc.) as well as executable documents which can be run to perform data analysis.

7.System Architecture

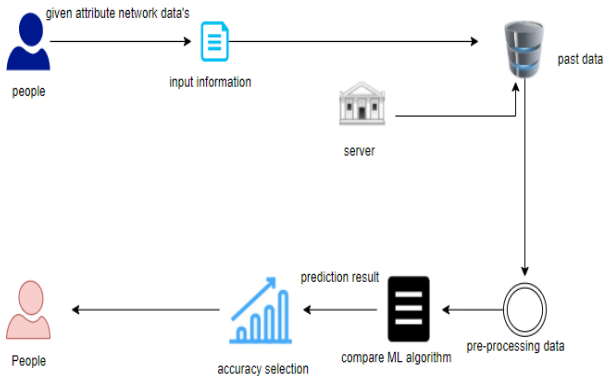

FIG-7: System Architecture

8. Module description

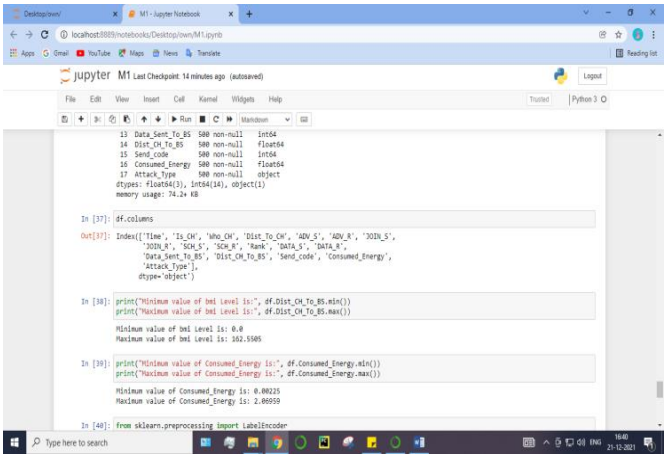

FIG-8.1:Data Validation and Pre-Processing

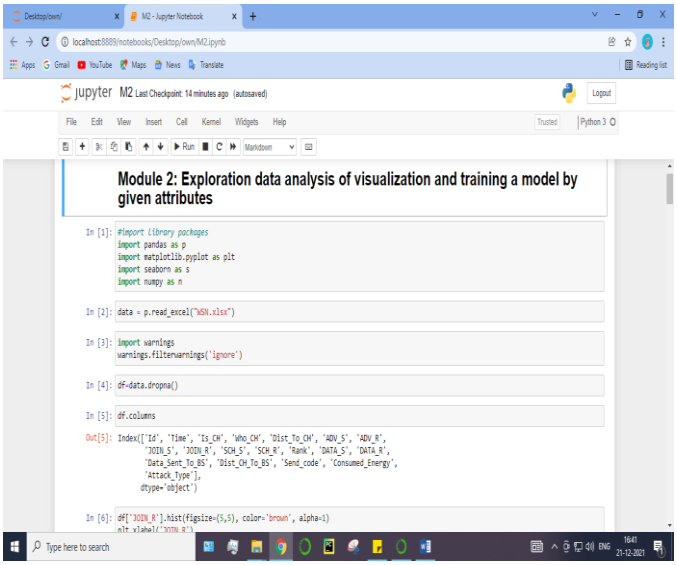

FIG-8.2:Exploration data analysis of visualization

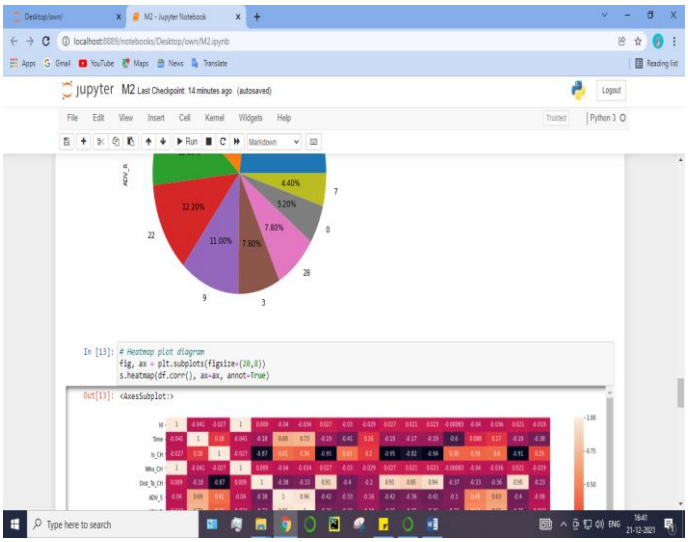

FIG-8.2.1

FIG-8.3:Attack Prediction

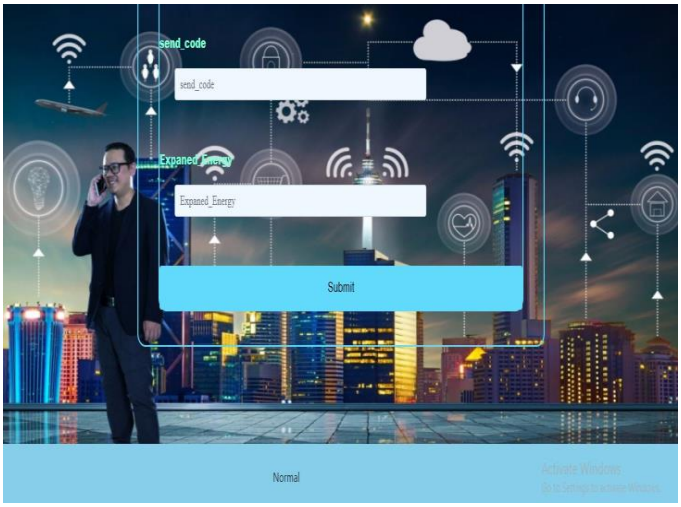

FIG-8.3.1

Logout

All Last

Time	Is_CH	who_CH	Dist_To_CH	ADV_S	ADV_R	JOIN_S	JOIN_R	SCH_S	SCH_R	Rank	DATA_S	DATA_R	Data_Sent_To_BS	dist_CH_To_BS	send_code	Expaned_Energy	Attack_type
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Normal
4	5	56	45	4	5	2	4	5	54	5	1	2	52	5	2	5	Normal
1	2	3	4	5	6	8	7	8	5	4	5	2	5	2	5	2	Black hole
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	8	Black hole
1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	74	8	Normal
159	753	654	852	159	753	654	852	159	654	755	623	5145	85	25	6	32	Flooding

FIG-8.3.2: Data Set

9. Conclusion

The analytical process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation. The best accuracy on public test set of higher accuracy score algorithm will be find out. The founded one is used in the application which can help to find the Wireless Sensor Network Attack Prediction.

The analytical process started from data cleaning and processing, missing value, exploratory analysis and finally model building and evaluation. The best accuracy on public test set of higher accuracy score algorithm will be find out. The founded one is used in the application which can help to find the Wireless Sensor Network Attack Prediction.

10. Reference

[1] S. Lee and T. Chung, "Data aggregation for wireless sensor networks using self-organizing map" in Artificial Intelligence and Simulation, Berlin, Germany:Springer-Verlag, vol. 3397, pp. 508-517, 2005.

[2] R. Arroyo-Valles, R. Alaiz-Rodriguez, A. Guerrero-Curieses and J. Cid-Sueiro, "Q-probabilistic routing in wireless sensor networks", *Proc. 3rd Int. Conf. Intell. Sensors Sensor Netw. Inf.*, pp. 1-6, 2007.

[3] C. Guestrin, P. Bodik, R. Thibaux, M. Paskin and S. Madden, "Distributed regression: An efficient framework for modeling sensor network data", *Proc. 3rd Int. Symp. Inf. Process. Sensor Netw.*, pp. 1-10, 2004.

[4] J. Barbancho, C. León, F. Molina and A. Barbancho, "A new QoS routing algorithm based on self-organizing maps for wireless sensor networks", *Telecommun. Syst.*, vol. 36, no. 1–3, pp. 73-83, Nov. 2007.

[5] J.-M. Kim, S.-H. Park, Y.-J. Han and T.-M. Chung, "CHEF: Cluster head election mechanism using fuzzy logic in wireless sensor networks", *Proc. 10th Int. Conf. Adv. Commun. Technol.*, vol. 1, pp. 654-659, 2008.

[6] W. B. Heinzelman, *Application-specific protocol architectures for wireless networks*, 2000.

[7] P. Zappi et al., "Activity recognition from on-body sensors: Accuracy-power trade-off by dynamic sensor selection" in Wireless Sensor Networks, Berlin, Germany: Springer-Verlag, pp. 17-33, 2008.