iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VII **Month of publication:** July 2021

DOI: https://doi.org/10.22214/ijraset.2021.37085

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Resolving Wireless Sensor Networks Issues using Machine Learning Techniques: A Review

Harshitha S[1], Ramesh B[2]
[1, 2]Department of Computer Science and Engineering, Malnad College of Engineering, Affiliated to VTU Hassan, Karnataka, India.

***Abstract:** Wireless Sensor Network (WSN) is the most encouraging advancements for some real-world applications due to its size, low cost, and easily deployable nature. Because of some factors like temperature, humidity, wind, pressure, battery life, etc., the performance of WSN will change dynamically, and therefore it requires depreciating dispensable redesign of the network. The conventional WSN is controlled by external programs, which makes it the networks hard to respond dynamically. In order to provide a quick response for dynamic changes, Machine learning (ML) techniques can be applied on WSN. Machine Learning is a technique that is self-instructed from experience and acts without human intervention or re-program. In this paper, Machine learning techniques for solving various issues in WSN are presented; we discussed machine learning techniques for anomaly, fault, and event detection. Finally, we presented the statistical analysis of Machine learning techniques used to solve the issues of WSN.*

***Keywords:** Wireless Sensor Networks, Machine Learning techniques, anomaly detection, fault detection, and event detection.*

I. INTRODUCTION

The computing processes can be made more efficient, reliable, and cost-effective by using Machine learning (ML) techniques. ML can produce models quickly and accurately for even more complex data. Supervised, unsupervised, semi-supervised, and reinforcement learning are the various types of ML techniques. The strength of ML lies in its capacity to give derived solutions through architecture to figure out how to improve its presentation.ML plays a vital role in various fields, including engineering, medical, and computing, because of its multidisciplinary nature. Advanced ML techniques have been applied to solve various issues in WSNs. Applying ML not just improves the presentation of WSNs and furthermore restricts the human mediation or re-program. Accessing a huge volume of information gathered by the sensors and pulling out the valuable data from the information is hard without ML.

A. Wireless Sensor Network (WSN)

WSN is mentioned to be a group of geographically scattered and devoted sensors for observing and recording the physical states of the environment and sorting the collected data at a focal area. The atmospheric conditions like temperature, sound, pollution levels, humidity, wind, and so on can be measured by WSN. A WSN can be defined as a spatially distributed minute device called sensor nodes that work collaboratively in order to transmit the data accumulated from the observed field through remote connections. The information gathered by way of the unique nodes is sent to a sink which both uses the records domestically or is attached to other networks; for example, the internet is connected through a gateway shown in Figure 1[1].

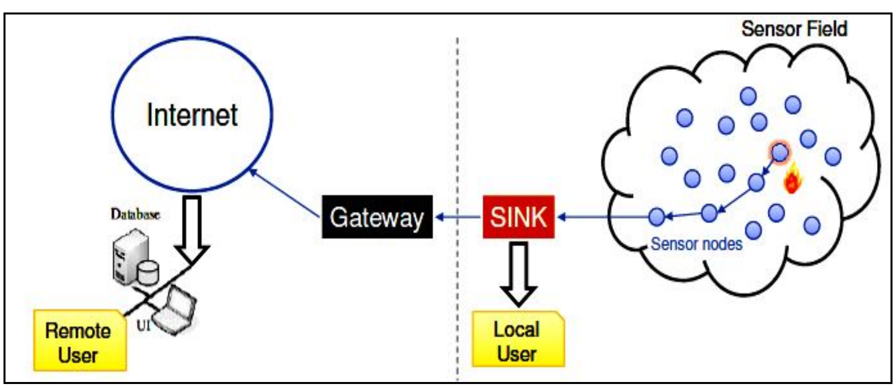


Fig 1: Wireless senor Network.

B. Machine Learning in WSN

Recent advances in ML have been applied to solve various issues in WSNs [1]. Applying ML not only improves the performance of WSNs and also limits the human intervention or re-program. Access vast amount of data collected by the sensors and extract useful information from the data is hard without ML. There are many WSN issues such as localization, coverage, and connective it, target tracking, routing, congestion control, anomaly detection, fault detection, event detection, data aggregation, synchronization, mobile sink, energy harvesting so on. All these issues can be solved using various machine learning techniques. In this paper majorly we have done a survey on anomaly detection, Fault detection, and event detection. The machine learning techniques and categories that are used to solve the WSN issues are discussed in this section. ML techniques have been categorized into supervised, unsupervised, semi-supervised, and reinforcement learning, as shown in Figure 2[2]. All the ML techniques are discussed by relating them to solve WSN issues.

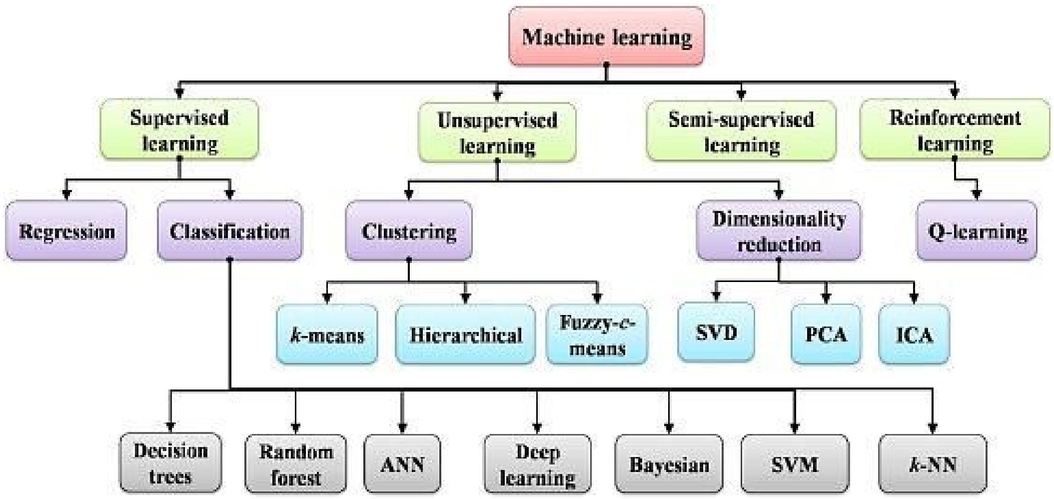


Fig2: Taxonomy of Machine learning techniques

1) *Supervised Learning:* One of the maximum critical facts processing approaches in ML is supervised learning. It provides a set of inputs and outputs(datasets and labels), and correlation between them is determined while training the system. We can find a function from an input *x* with the best estimation of output *y* (f: $x \rightarrow y$)at the end of training. Localization, coverage problems, anomaly and fault detection, routing, data aggregation, synchronization, congestion control, target tracking, event detection, and energy harvesting are the issues solved using supervised learning. Supervised learning is divided into regression and classification. Logic-based are decision tree and random forest, perceptron-based is ANN and deep learning, statistical learning is Bayesian, SVM, and instance-based are k-NN algorithms are the various subdivisions of classification.

a) *Regression:* Regression is a supervised gaining knowledge approach, and it will expect a few continuous values (Y*)* based on a given multiple set of features (X*)*. The variables in the regression model are uninterrupted or measurable. It is an extremely basic ML approach and predicts exact outcomes with the least blunders.

b) *Decision Trees:* Decision trees (DT) are a category of supervised ML approaches for classification. It is based on a set of if-then rules to increase understandability. The decision node and Leaf node are the two types of nodes in a Decision tree where the leaf node gives the final result, and the decision node helps to make the selection between different choices. A class label or a target is predicted by designing a training model based on choice-based rules obtained from training data.

c) *Random-forest:* Random-forest (RF) algorithm is a supervised ML technique. It consists of the number of decision trees like a forest where all the trees can help to predict and make the classification. Creating random forest and predicting the results are the two phases of RF algorithm. RF turns out productively for the bigger datasets and heterogeneous information. This methodology precisely predicts the missing qualities.

d) *Artificial Neural Networks:* An artificial neural network (ANN) is a supervised ML technique. In view of the model of a human neuron for arranging the information ANN is designed. ANN is associated with an immense number of neurons (handling units) that cycle data and produce precise outcomes. ANN normally works on layers, these layers associated with hubs and every hub related with an active function. Input, hidden and output are the three layers. ANN characterizes perplexing and non-straight informational collections effectively, and there is no limitation for the data sources like other classification strategies. Even though ANN has a higher computation requirement, it is still used in many real-time applications of WSN.

e) *Deep Learning:* ANN is subdivided into Deep learning, which is also a supervised ML approach used for classification. Deep learning approaches are the information learning portrayal techniques with multiple layer presentation. It makes with basic non-straight modules that change the presentation from lower layer to higher layer to accomplish the best results. The vital advantages of Deep learning are separating high-level attributes from the data, work with or without name labels, and it very well may be prepared to satisfy numerous targets.

f) *Support vector-machine:* Support vector machine (SVM) is a supervised ML classifier that tracks down an ideal hyperplane to classify the information. SVM plays out the best order utilizing hyperplane and arrange singular perception. The greater part of the preparation information is repetitive once a limit is set up and a bunch of focuses assists with recognizing the limit. The focuses which are utilized to discover the limit are called support vectors. SVM gives the best characterization from a given classification of information.

g) *Bayesian:* Based on statistical learning approaches, a supervised ML algorithm is designed, which is called as Bayesian. By using several statistical methods like the Chi-square test, Bayesian learning discovers the connections among the datasets by learning conditional independence. Bayesian learning strategies can be used to solve several WSNs problems to improve network efficiency.

h) *K -Nearest Neighbor:* The most direct and lethargic, instance-based learning method in regression and classification is K - Nearest Neighbor (k -NN). The k - nearest training set consider as a contribution from the component space. K-NN generally characterizes based on the distance between specified training samples and the test sample. Euclidean distance, Hamming distance, Manhattan distance, Minkowski distance are the various distance functions used in K-NN. The intricacy of the k - NN calculation relies upon the size of the input dataset and ideal execution if a similar size of the information. This methodology tracks down the conceivable missing qualities from the element space and, furthermore, lessens the dimensionality.

2) *Unsupervised Learning:* The inputs are not associated with labels in unsupervised learning. In other words, there are no class labels, but still, the model tries to extract the relationships from the data. Classifying the set of comparable patterns into clusters, dimensionality discount, and anomaly detection from the data are the various approaches of unsupervised learning.

a) *k-means Clustering:* The k -means algorithm effectively shapes a specific number of clusters from a given dataset. At first, the k number of arbitrary areas is considered, and all the remaining points are associated with the nearest centers. Further, once the clusters are formed by covering all the points from the dataset, a new centroid from each cluster is re-calculated. In each iteration, the centroid of the cluster changes, and repeat the algorithm until there will be no more changes in the centroid of all clusters. To find optimal cluster heads for routing the data towards to base station in WSN, k -means clustering is the simplest clustering which will be more useful. This methodology is likewise helpful to track down the efficient rendezvous points for the mobile sink.

b) *Hierarchical Clustering:* Grouping similar objects into clusters that have a pre-decided top-down or bottom-up order is called the Hierarchical clustering technique. Divisive clustering is another name given to top-down hierarchical clustering; in this clustering, a huge single partition split recursively until one cluster for every perception. Agglomerative clustering is another name given to bottom-up hierarchical clustering; in this approach, every perception appoints to its cluster dependent on density function.

c) *Fuzzy- c -means clustering:* Fuzzy- c -mean (FCM) clustering is also named as soft clustering using fuzzy set theory, where it assigns the perception to one or more clusters. In this method, based on the similarity measurements such as the intensity, distance, or connectivity clusters are identified. Depending upon the applications or datasets, the algorithms may consider at least one similarity measure. The algorithm emphasizes the clusters to track down the optimal cluster centers. FCM produces the optimal clustering when contrasted with k - means for the over-lapped datasets.

d) *Singular Value Decomposition:* In order to reduce the dimensionality, one of the matrix factorization methods is used, and it is called Singular value decomposition (SVD). Representing a matrix into a product of matrices is called the Matrix factorization method. SVD can be utilized proficiently for decreasing the data dimensionality of the given element space. SVD is utilized in WSNs to resolve different issues like routing and data aggregation.

e) *Principle Component Analysis:* In order to perform dimensionality reduction, a multivariate analysis feature extraction method called Principle component analysis (PCA) is used. The PCA joins all the data and drops the least need data from the component space to lessen the dimensionality. The yield of PCA is a linear combination of noticed factors (principal components).In PCA, some-time used to distinguish inconsistencies from the information, just as in regression. Sensors continuously gather information from the environments and transmit it to the base station in WSN. Applying PCA can reduce the dimensionality of the data either at sensor level or at cluster head level to reduce the communication overheads in WSN.

f) Independent Component Analysis: Independent component analysis (ICA) tracks down another reason for information representation and decomposes multivariate perceptions into added substance subcomponents. Here the subcomponents are non-Gaussian perceptions. ICA is a more remarkable method than PCA, at the end of the day, its an extension of the PCA. ICA will eliminate the higher-order dependencies, while PCA couldn't do.

3) Semi-supervised Learning: The vast majority of this present reality application's information is labeled and unlabeled. The supervised learning algorithms work productively on the labeled data, and unsupervised learning works effectively on unlabeled data. The semi-supervised learning is acquainted with work on the information with the blend of both labeled and unlabeled. Predicting the labels on unlabeled data in the training set and to predicting the labels on future test data sets are two distinct goals of semi-supervised learning. Concerning these goals, semi-supervised learning is categorized into two types: Transductive learning and inductive semi-supervised learning. The inductive semi-supervised learning just learns a function, whereas the Transductive learning is used to predict the exact labels for a given unlabeled dataset. f : X->Y so that f expected to be a good predictor on future data. Semi-supervised learning is used in many real-time applications

4) Reinforcement Learning: Reinforcement learning (RL) algorithm constantly learns by inter-acting with the environment and assembles data to make certain moves. RL increases the performance by getting the ideal outcome from the environment.

II. LITERATURE SURVEY

There are many issues in WSN. A lot of research is carried on various WSN issues. In each sub-section, we discuss the advantages of selecting an ML technique to address the issues in WSNs and also present the features of existing approaches in the tabular form. In this paper, we studied and analyzed anomaly, fault, and event detection in WSN.

A. Anomaly detection

A differed variation observed significantly from information readings is called an anomaly. In WSNs, misconduct is distinguished either in estimating the sensor information or in traffic-related properties. A large portion of the applications in WSNs, sensors continuously assembling the information from the environment and send it to the base station through relay hubs. Because of abnormal attacks, there is data loss during data transmission. So there is a need for protecting sensing data. So, while sharing the information between the sensor nodes, anomaly detection in WSNs minimizes the communication overhead.

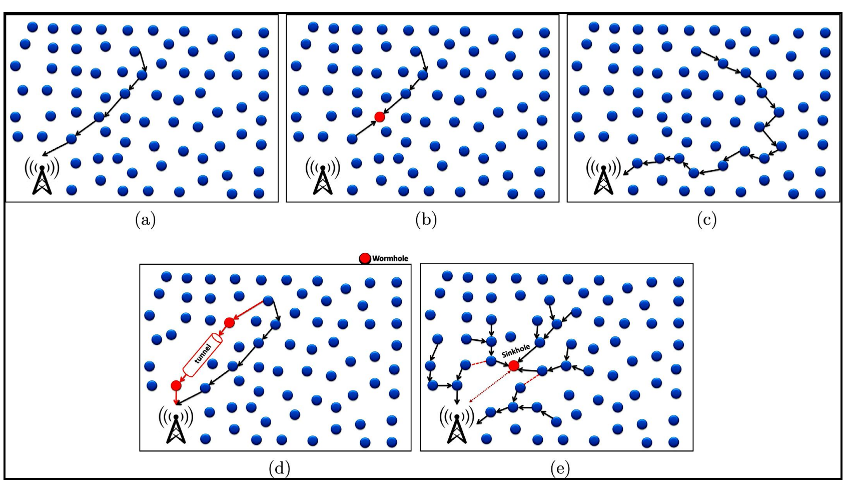


Fig 3 Anomaly detection in WSNs (a) normal flow (b) black hole attack (c) misdirection attack (d) wormhole attack (e) sinkhole attack.

Blackhole attack, misdirection attack, wormhole attack, sinkhole attack, and hybrid anomaly are some of the possible attacks in WSNs, as shown in Figure 3. In a black hole attack, a block gets the packets as opposed to sending them to the base station. In a misdirection attack, the attackers send the packets to distinct nodes, as opposes to sending it to its neighbor nodes. This may lead to a longer route which reduces the throughput. In a wormhole (WH) attack, between two distinct nodes, a WH tunnel is formed and misunderstands that they are extremely close. The number of packets from the network are bypassed or attracted by this WH and attackers can perform the manipulations. The sinkhole is a particular node that advertises ideal route to its neighbor nodes to the base station. The data has tampered and the network is damaged in a sinkhole attack. To detect anomalies in WSNs, Protecting from various assaults and misapprehensions, various ML approaches have been used.

A hybrid anomaly is the combination of diverse assaults, therefore detecting the node which effects and type of anomaly are happening. Employing clustering algorithms minimizes the complexity and conversation overhead for the trouble. For detecting the anomalies in non-stationary environments, ML techniques assure to deal with faults, assaults, and outliers in WSNs. For online anomaly detection, it is viable to adjust the parameters dynamically the use of the historic statistics.

[3] Efficient Hybrid anomaly detection using K – means clustering is proposed .Hybrid anomaly is the type which contains the various types of attacker nodes such as black hole , worm hole so on. In this situation it would be very difficult to save the network and increase the lifetime of the network. So Mohammad Wazid et al, proposed the secure intrusion detection approach using K-means clustering. Once the detected intrusions of network activities are compared against detected patterns , it has been evaluated using WSN dataset created using opnet modeller which has got the features like end-to-end delay, sent traffic and received traffic. This technique has the ability to detect two types of malicious nodes: black hole and misdirection nodes with 98.6 % detection rate and 1.2 % false positive rate.

[4] Author discussed about the feasibility of various techniques to detect anomaly in harsh environments. They presented complete analysis of one-class SVM for outlier and event detection in WSN. These include hyper-plane, hyper-sphere, quarter-sphere and hyper-ellipsoid. From the analysis it has been observed that quarter-sphere formulations are most suitable for outlier detection in WSNs because of their less computation and transmission complexity and high efficiency.

[5] Authors have designed a new architecture called intrusion detection systems (IDS) makes use of selection tree class method to make a certain high detection rate. This approach is specifically utilized in environmental monitoring.

[6] H S Emadi et al,.proposed a novel anomaly detection algorithm in WSN using DBSCAN and SVM. Here the features from temperature, humidity and voltage are taken from the network traffic and analysed using the DBSCAN algorithm. The low-density points in the regions are detected as an anomaly using this algorithm. The SVM is trained using the normal data.At last anomalies are removed from network traffic. The standard and general data set of Intel Berkeley Research lab (IRLB) were used for evaluation by the algorithm proposed in this paper. In this paper, Using the soft computing methods, very simple implementation and improved accuracy detection rate by analyzing all the features simultaneously are the advantages added to this work.

[7] Z.Feng et al,. discussed an anomaly detection method in WSN by proposing a improved method of support vector data description (SVDD). Their main aim was to reduce the computational complexity in training and testing phases and apply the method to detect the anomaly in the data node of WSN. The strategy for reducing training set and Sequential minimal optimization(SMO) is used to improve the training speed. They have used standard datasets like University of California Irvine (UCI) datasets. The proposed new SVDD is evaluated using the real Intel Berkeley Research Lab (IBRL) and the labeled WSNs datasets.

[8]Tao Ma et al,. proposed an intrusion detection system to detect malicious attacks in network traffic. They gave a new approach by combining spectral clustering(SC) and Deep Neural network[DNN] which is called SCDNN. Based on the sample similarity the dataset is divided into k subsets as in SC and again based on the similarity features the distance between data points in testing and training set is measured. Later this is trained in deep neural network algorithm. Six KDD-Cup99, NSL-KDD datasets and a sensor network dataset were used to test the performance of the model. Results proved that the SCDNN classifier performs better than backpropagation neural network, SVM,Random forest and Bayes tree models in terms of detection accuracy.

Table I: Machine Learning Based Anomaly Detection Techniques In WSN

Author	Machine Learning Algorithm	Anamoly	Data sets	Results
[3]	K – means	Black hole and misdirection nodes	Genearated data from Opnet simulator	98.6
[4]	SVM	Harsh environments	Simulated data	97.5
[5]	Decision Tree	Sink hole	AODV routing protocol simulated	99.5
[6]	DBSCAN and SVM	Low density points	Intel Berkeley Research lab	97
[7]	Support vector data description	data node	University of California Irvine	Improved Training time and testing speed.
[8]	SCDNN	malicious attacks	KDD-Cup99, NSL-KDD	92.88

Using ML for anomaly detection in WSNs significantly improves detection rate as compared to other approaches. Employing clustering algorithms minimizes the complexity and communication overhead for the problem. ML techniques are use full for detecting the anomalies in non-stationary environments, and ML approaches guarantee to handle faults, attacks, and outliers in WSNs. For online anomaly detection, it is possible to adjust the parameters dynamically using the historical information. Table I, summarizes the ML approaches for detecting an anomaly in WSNs.

B. Fault Detection

The sensor nodes deployed in various locations may sometimes be not reachable and unkind due to harsh environmental factors. Hence, faults may occur in WSNs such as transmission failures, battery failures, hardware and software failures, ineffective base stations, or changes in topology. The resource limitations, difficulty in distinguishing normal and faulty nodes, Environmental factors, changes in deployment have to lead to face number of challenges in detecting the faults in WSN. ML approaches increase the accuracy rate in detecting faults because of their quick detection and prediction of faults

[9] proposed a centralized solution to detect faults in WSN using a Deep neural network. The main aim of this paper was to enhance the detection accuracy of faults and power consumption. Experiments were conducted by simulation in Matlab, and they took the KDD99 database with 41 features. A very high detection accuracy rate is achieved in the results. They also overcome the challenges faced while training the distributed layers of DNN.

[10] a novel intrusion detection system using K-nearest neighbor classification in WSN is proposed. This classifier classifies the normal nodes and faulty nodes accurately by analyzing the abnormal behavior of faulty nodes.

[11] authors have taken the datasets from various observation vectors in three different instances, and then the datasets were induced with two faults. The different classifiers were applied, such as SVM, CNN,RF so on, and made the comparison study on the accuracy of fault detection rate and concluded that Random forest (RF) classifier gave the optimal results.

[12] Yong Cheng et al. proposed the new fault detection mechanism based on support vector regression. A fault prediction model is built using the support vector regression from the unessential information about climatic elements collected by multi-sensors. Simulations conducted in Matlab prove that the proposed algorithm gives an accuracy of 87% when the sensor fault probability is WSN is 40%.The false alarm rate(FAR) ratio is below 7%. So this algorithm provides high detection accuracy and a low false alarm ratio.

An error prediction method has been developed in [13] the usage of SVM and cuckoo seek algorithms. In this approach, SVM used to expect the mistakes in the sensor dynamically; however, it relies upon the parameter comparisons. To optimize the important thing parameter, this technique followed the cuckoo seek set of rules to keep away from the nearby minimum value.

Table Ii: Machine Learning Based Fault Detection Techniques In WSN

Author	Machine Learning Algorithm	Fault	Data sets	Results
[9]	DNN	Reducing the fusion center's calculating load and saving the WSN's transmitting power consumption	KDD99 database	99
[10]	KNN classification	Abnormal nodes	Network simulated data	99.0 and FAR is 1.5
[11]	Random forest	Spike and data loss fault	Real world data	93.15
[12]	Support vector regression	Residual sequence collected from muti sensors	Network simulated data	87 and False Alarm Ratio is below 7
[13]	CS-SVM	Dynamically finds the mistakes in sensors	Generated datasets	98

By using ML algorithm, Brisk detection and categorization of faults are possible. So it improves accuracy in detecting faults. ML-based fault detection techniques for WSNs are shown in Table II.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

C. Event detection

In WSNs, sensor hubs constantly screen the climate, notice an event and method locally and select some specific choices. The component of identifying the event or misbehavior from the information is referred to as event detection. The necessities to fulfill the event recognition are synchronization, low false rate, and high true detection rate. Event detection is challenging due to sensors with less power, memory, and computational resources. Therefore ML-based approaches have the prospective to overcome event detection problems. Even in the complex forms of sensor data ML can easily detect the event and also improves the packet delivery ratio with efficient duty cycles.

[14] the authors have used a sophisticated regression model to enhance the event detection accuracy and K-NN was applied to extract essential data from the sensors effectively. They mainly concentrated on detecting three different kind of event for various WSN applications. Mainly divided the data into two datasets. One is the historical dataset from which they learnt estimation models and other parameters. Another was the test dataset which gives the online measurements and used for evaluation. The datasets for three different contexts were obtained from PhysioNet MIMIC II Waveform Database, NIST website, Harvard Sensor Networks Lab. So this paper is focused on efficiently detecting malicious data injections in event detection in WSN.

[15] author proposes a new event-based kNN query scheme for distributed sensor systems using fuzzy sets. K-NN query is the best method to extract useful information from distributed sensor devices but they are costly in terms of transmission and time overhead. So linguistic e-KNN information is used in this paper which is very useful in energy consumption. Experiments are conducted on real life datasets and it has been proved that e-KNN method works best than the traditional methods. It guarantees the low communication cost and response time with increased accuracy.

[16]M.M.Hassanetal,.proposed a human activity recognition system using smartphone sensors through deep learning approach. This system is very helpful in applications like smart home health care to monitor and rehabilitate the patients conditions.In this paper , a smartphone inertial-based sensor approach is used where effective features are extracted raw data. Mean, median, autoregressive coefficients etc are selected as features. Features are still processed by applying kernel principal component analysis(KPCA) and Linear discriminant analysis(LDA) to make them robust and at last they are trained using Deep Belief network. This approach is also compared with conventional recognition approaches, where it outperformed well.

Table Iii: Machine Learning Based Event Detection Techniques In WSN

Author	Machine Learning Algorithm	Event	Data sets	Results
[14]	Regression model and KNN	Malicious data injections in event detection	PhysioNet MIMIC II Waveform Database, NIST website, Harvard Sensor Networks Lab	97
[15]	e-KNN using fuzzy sets	Extract information of interest from distributed sensing devices with energy efficiency	Simulated and real life datasets	98.7
[16]	DBN,KPCA and LDA	Human activity recognition system	Simulated data	96

Machine learning approaches are very useful to detect an event from the complex forms of sensor data. ML also improves the packet delivery ratio by achieving efficient duty cycles. Table III shows some of the work on Machine Learning-based Event detection techniques in WSN.

III.STATISTICAL ANALYSIS OF ML-BASED ALGORITHMS FOR WSNS

The statistical charts give an overview of recent research on ML-based algorithms for WSNs is as shown in Figure 4[2]. From Figure 4(a) and Figure 4(b), we notice that most researchers focus on data aggregation, localization, routing, anomaly and fault detection, event detection, coverage, and connectivity. Figure 4(a) shows the percentage of research work investigated in solving WSN issues using machine learning techniques. Figure 4(b) shows the year-wise research conducted on WSN issues. Figure 4(c) estimates that most of the WSNs issues were solved using supervised learning algorithms, and Figure 4 (d) that most of the WSNs issues are solved by using the Bayesian learning approach.

From the literature survey, it has been identified that still a lot of research is in active mode on using ML techniques in WSN. Anomaly detection is one of the major issues in WSN. Anomalies in the network may lead to communication overhead, delay in transmission and sometimes due to anomalies, data is routed to the wrong destinations. So detecting the anomalies accurately, reducing false alarm rate and improving the security in the network is of major concern in today's fast-growing use of the internet. Many Machine learning techniques are used to train the data in the network to detect malicious attacks by designing various intrusion detection systems. As discussed in the literature survey, SVM classifiers, Random Forest classifiers, SVDD, Artificial neural networks, Deep neural networks can be used to detect anomalies. Fault detection is also an issue in WSN where it may degrade the performance due to hardware and software failures, battery failures, transmission failure so on. It is very important to detect abnormal and faulty nodes. This can be achieved through ML techniques such as K-nearest neighbor, DNN, SVM, CNN, RF. Event detection is the one where we detect misbehavior from the data. Here also high detection rate and low false alarm rate and synchronization are the requirements to detect the events. Sophisticated regression models, K-NN, e-KNN, kernel PCA, LDA, Deep belief network(DBN), are the ML techniques that are used in event detection.

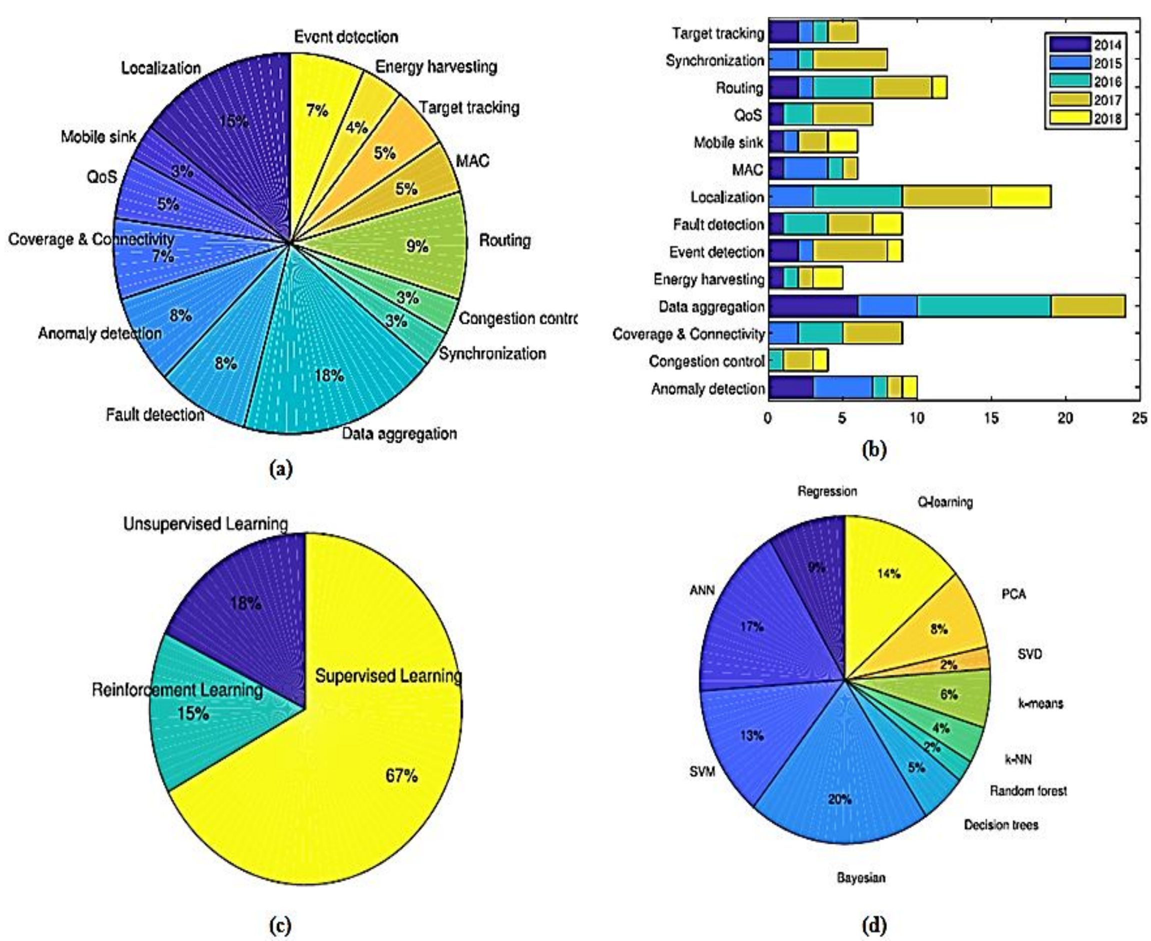


Fig 4 . Statistical charts (a) Issues in WSNs addressed by ML algorithms (2014–March 2018) (b) Year-wise research articles (c) Classification of ML algorithms for WSNs (d) ML-based algorithms for WSNs.

IV. CONCLUSIONS

The survey on anamoly, fault, and event detection issues of wireless sensor networks are discussed in this paper and also identified more than one Machine Learning technique to overcome these issues. From the statistical analysis of previous research work, it has been identified that many WSN issues can be solved using a machine learning algorithm. The supervised machine learning algorithm shows better performance over the unsupervised machine learning algorithm. Due to the increase in usage of multimedia from the past decade, further investigation on WSN issues can be explored using machine learning techniques.

REFERENCES

[1] PriyankaRawat , Kamal Deep Singh ,HakimaChaouchi · Jean Marie BonninWireless sensor networks: A survey on recent developments and potential synergies,April 2013,The Journal of Supercomputing 68(1) DOI: 10.1007/s11227-013-1021-9.

[2] D. Praveen Kumar, TarachandAmgoth , Chandra SekharaRaoAnnavarapu,,Machine learning algorithms for wireless sensor networks: A survey, Information FusionVolume 49, September 2019, Pages 1-25

[3] Mohammad Wazid & Ashok Kumar DasAn Efficient Hybrid Anomaly Detection Scheme Using K-Means Clustering for Wireless Sensor Networks,. Wireless Personal communications volume 90, pages1971–2000 (2016)

[4] N. Shahid , I.H. Naqvi , S.B. Qaisar , One-class support vector machines: analysis of outlier detection for wireless sensor networks in harsh environments, Artif. Intell. Rev. 43 (4) (2015) 515–563

[5] Enhancing Intrusion Detection in Wireless Sensor Networks through Decision Trees AlessiaGarofalo, Cesario Di Sarno, and Valerio Formicola ,Department of Technology, University of Naples Parthenope, Naples, Italy

[6] H.S. Emadi , S.M. Mazinani , A novel anomaly detection algorithm using DBSCAN and SVM in wireless sensor networks, Wireless Personal Commun. 98 (2) (2018) 2025–2035

[7] Z. Feng , J. Fu , D. Du , F. Li , S. Sun , A new approach of anomaly detection in wireless sensor networks using support vector data description, Int. J. Distrib. Sens. Netw. 13 (1) (2017) 1–14 .

[8] T. Ma , F. Wang , J. Cheng , Y. Yu , X. Chen , A hybrid spectral clustering and deep neu- ral network ensemble algorithm for intrusion detection in sensor networks, Sensors 16 (10) (2016) 1–23 .

[9] C. Li , X. Xie , Y. Huang , H. Wang , C. Niu , Distributed data mining based on deep neu- ral network for wireless sensor network, Int. J. Distrib. Sens. Netw. 11 (7) (2015) 1–7 .

[10] W. Li , P. Yi , Y. Wu , L. Pan , J. Li , A new intrusion detection system based on KNN classification algorithm in wireless sensor network, J. Electr. Comput. Eng. 2014 (2014) 1–9 .

[11] Fault Detection in Wireless Sensor Networks through the Random Forest Classifier ZainibNoshad,Nadeem Javaid,1,Tanzila Saba,ZahidWadud,MuhammadQaiserSaleem, Mohammad EidAlzahrani,and Osama E. Sheta, Journal ListSensors (Basel)v.19(7); 2019 Apr PMC6480196

[12] Distributed Fault Detection for Wireless Sensor Networks Based on Support Vector Regression Yong Cheng, Qiuyue Liu, Jun Wang, Shaohua Wan , and Tariq Umer, October 2018Wireless Communications and Mobile Computing 2018(5):1-8,Volume 2018 |Article ID 4349795 | https://doi.org/10.1155/2018/4349795

[13] M. Jiang , J. Luo , D. Jiang , J. Xiong , H. Song , J. Shen , A cuckoo search-support vector machine model for predicting dynamic measurement errors of sensors, IEEE Access 4 (2016) 5030–5037

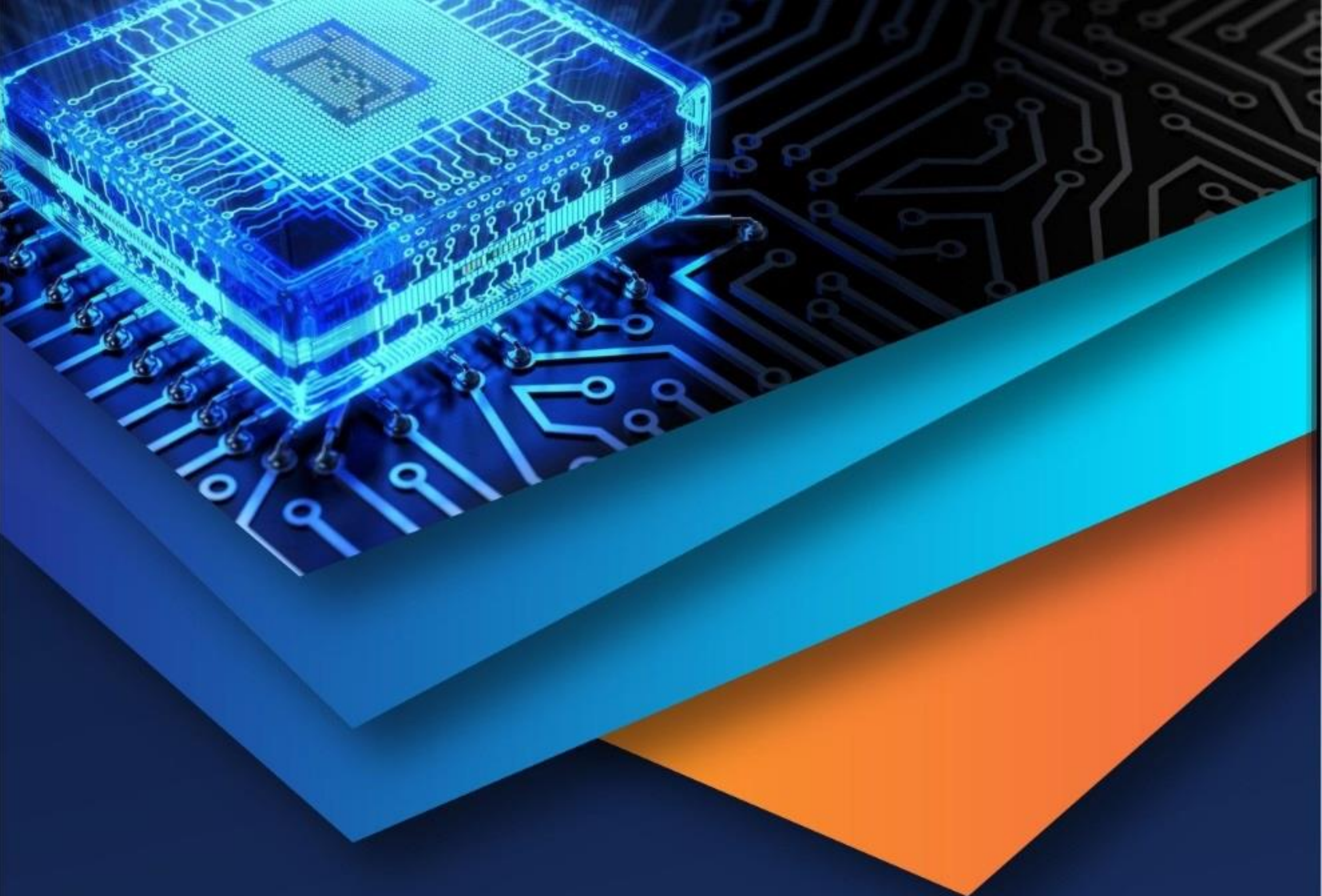
[14] V.P. Illiano , E.C. Lupu , Detecting malicious data injections in event detection wire- less sensor networks, IEEE Trans. Netw. Serv. Manage. 12 (3) (2015) 496–510 .

[15] Y. Li , H. Chen , M. Lv , Y. Li , Event-based k-nearest neighbors query processing over distributed sensory data using fuzzy sets, Soft Comput. (2017) 1–13

[16] M.M. Hassan , M.Z. Uddin , A. Mohamed , A. Almogren , A robust human activity recognition system using smartphone sensors and deep learning, Future Gener. Comput. Syst. 81 (2018) 307–313 .

[17] Adaptive image compression technique for wireless sensornetworksqMohsenNasri et.al, Computers & Electrical EngineeringVolume 37, Issue 5, September 2011, Pages 798-810

[18] Jayanthi Ramasamy , John Singh Kumaresan Image Encryption and Cluster Based Framework for Secured Image Transmission in Wireless Sensor Networks. Wireless Personal Communications (2020) 112:1355–1368.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY