Wireless Body Area Network Sensor Faults and Anomalous Data Detection and Classification using Machine Learning

Sumit Kumar Nagdeo and Judhistir Mahapatro
Department of Computer Science and Engineering
National Institute of Technology, Rourkela
Odisha, India - 769008
Email: {217CS3315, mahapatroj}@nitrkl.ac.in

***Abstract*—Sensor Networks are very much vulnerable and prone to faults and external attacks. Sensor networks used for Healthcare Monitoring are termed as Wireless Body Area Networks (WBAN), which is used for collecting various vital physiological parameters of patients from remote locations. However, WBAN sensors are prone to failures because of noise, hardware misplacement, patient‘s sweating. Sensed data from these sensors are sent from the Local Processing Unit to Medical Professionals. It would be very difficult for the Medical Professionals to diagnose correctly if the sensed data from these sensors are faulty or effected by the malicious third party. At times, even faulty data might lead to misdiagnosis or death of a patient. It motivated us to address this challenge by proposing a Machine Learning Paradigm to distinguish this anomalous data from the genuine sensed data. Firstly, we classify the health parameters as normal records or abnormal record. After the classification, we propose to apply regression technique for identifying the anomalous data and actual critical data. We use real patient‘s vital physiological parameters for validating the robustness and reliability of our proposed approach.**

***Index Terms*—Sensor Networks, Wireless Body Area Network, Regression Model, True Positive Rate, False Positive Rate.**

I. Introduction

With increasing population and increase in elderly patients, it has led to increment in the medicinal services expenses and deficiency of healthcare experts [1], which leads to increase in the popularity of remote healthcare monitoring of the patients.

Healthcare Monitoring System consist of various medical wireless sensors, referred as WBAN. WBAN represents Wireless Body Area Network, which transmit various health related parameters such as Blood Pressure (BP), Pulse, Respiration Rate (RR), Heart Rate (HR), ElectroCardioGram (ECG), Oxygen Saturation (SpO2), etc. using a Local Processing Unit (LPU) to healthcare professionals to monitor the patient remotely. LPUs(Smartphone/Tablet) are needed to send the collected data from sensors to the caregivers as they have as better computation and battery power, better transmission range and bandwidth. It is also used to receive diagnosis or any other instructions from caregivers. LPUs must be robust and reliable in monitoring real time data, and raising the medical alarms for caregivers during medical emergency [2].

Emergency condition for a patients arise when irregularities in the sensed data is observed multiple times. For example, normal range for HR lies between 60 and 90, if the HR value is above 90 or below 60, then it can be considered as the irregularity in the HR. And, when these irregularities are repeated multiple times, then that will be the emergency condition, and the system must raise an alarm for caregiver.

WBAN sensors have many advantages when it comes to Healthcare Monitoring System, it reduces the cost and improves the monitoring of healthcare by giving the caregivers opportunity to monitors patients remotely and constantly. WBAN sensors can increase the chance of discovering diseases at early stage. However, its low computing power and smaller battery life leads to poor reliability and vulnerable for security attacks after deployment. Sensor readings can both be inaccurate and unreliable [3] [4] [5], resulting from various hardware as well as software constraints. Every single reading sensed by the sensors are prone to noise, interference, misplacement of sensors, sweating patients and external hacks such as data injection and modification. This vulnerability of the WBAN sensors might result in false alarm raising for the caregiver, which makes the entire system very unreliable.

In our paper, we propose a mechanism based on the paradigm [6], which intelligently distinguishes whether the sensed data is due to irregularity in the patient‘s health or the sensor failure. We use Artificial Neural Network(ANN), to detect abnormal records from the dataset, then for every abnormal record identified, we determine to find predicted value for the records to find whether the sensed record is anomaly or not. However, physiological parameters are mostly co-dependent, and any changes occur in one parameter at least affect two or more parameters at the same time. For example, In asthma, the respiration rate and heart rate or pulse increases simultaneously. It has also been observed the HR and Blood Pressure(BP) also increases or decrease at the same time for cardiac-related disease.

The remaining of the paper is organised as follows: In Section II, we have presented the previously established methods which detect anomalous data out of Medical Sensors. In Section III, we have briefly presented our modified approach to the already presented solution used in our Healthcare Monitoring System. In Section IV, experimental results of our proposed solution are shows, where we tested our solution with real patient data. At last in section V, we have concluded the paper.

II. Related Work

Healthcare Monitoring System is getting prevalent day by day as a result of its productive and remote checking which saves both time and money for the patients as well as healthcare professionals. Different Healthcare monitoring systems have been proposed, created and are currently in use, for example,

- MEDiSN [3] and CodeBlue [7], which are used to monitor HR, ECG and SpO2.
- LifeGuard [8], which is used to monitor ECG, breath, beat oximeter & BP.
- AlarmNet [9] & Medical MoteCare [10], which are used to monitor physiological parameters such as heart beat & SpO2 & natural parameters such as temperature & light.
- Vital Jacket [11], which monitors ECG & HR

Study of Healthcare monitoring application using WBANs are available in [12] [13]. All of these applications shows similar type of challenges such as limited battery life, faulty sensors, third party infection, inefficiency due to patient‘s sweating. As Patient‘s health and diagnosis highly depends upon the effective working of these sensors, any type of failure or irregularity in the physiological data sent from these sensors is not at all affordable for the patients, many researchers have already proposed different methods of autonomous fault detection and classification for WSN and WBAN. One of such method is cluster based algorithm, which is used to classify outliers from faulty or malicious sensors, proposed by the authors in [14]. But the proposed method is impractical for medical sensors, as we Healthcare Monitoring System does not contain redundant medical sensors for monitoring the physiological parameters.

From many years now, SVM classification is the most popular classification machine learning model used to classify data and is often found to be the most optimal solution for most of the problems. Various modern SVM based approaches are proposed for Anomalous Data detection and classification in WSNs [5] [15] [16].

A effective approach for Anomaly Detection using J48 and Linear Regression is proposed in [6]. In their approach, they first classify the Sensed data from the Medical Sensors into normal or abnormal using J48 Decision Tree model, and after the classification of Abnormal Data, using Linear Regression, a value is predicted for that abnormal data using the other dependable parameters to check whether the sensed data is anomaly or not, by measuring the error, and if error found is greater than the pre-specified threshold, alarm will raise for the caregiver to monitor the emergency condition.

In this paper, we tend to enhance the proposed approach by changing the initially used J48 Decision Tree model with Artificial Neural Network (ANN), which results greater accuracy for classification of the abnormalities, then the Ensemble LinReg model is used to distinguish between the anomalous data from the detected abnormalities. Instances containing anomalous data will be discarded by the system to reduce the false alarm, which will make the system more reliable and efficient for use. In comparison to J48 Decision Tree model, ANN uses Feature Extraction in the pre-processing which enhances the accuracy for the classification. We are also comparing our results with using SVM and Linear Regression [17], which states out of all the three proposed model, our proposed approach of ANN with Ensemble LinReg yields better results with maximum accuracy, least error and reduced false positive rate.

III. Methodology

We consider the following architecture for our Healthcare Monitoring System, in which WBAN sensors are connected to body of the patient to monitor & transmit physiological health parameters, as shown in Figure 1.

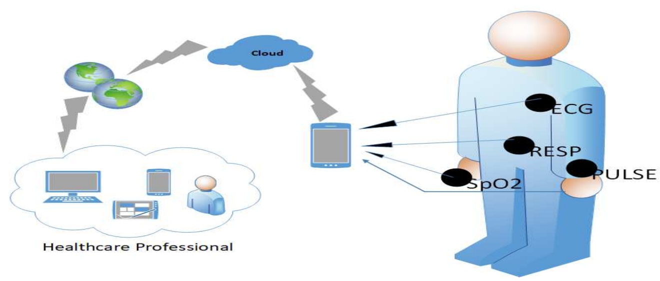

Fig. 1. Healthcare Monitoring System Architecture

These sensed data from WBAN sensors is received by LPU for analysis and raising alarm for Caregivers, if needed. Our proposed approach will be deployed in the LPU, as LPU has better computation capability with better power backup than the WBAN sensors. In LPU, sensed data will be analysed to classify and detect the anomalous data from genuine data, and alert healthcare professionals for any emergency condition. These sensed data are sent via network to Healthcare Professionals for monitoring and diagnosis of the patients.

The sensed physiological parameters from WBAN sensors are denoted by matrix $Z = (Z_{ij})$ where i is the instance and j is the sensed physiological parameter. Lets assume, there are m instances and n physiological parameters sensed through WBAN sensors. Then, the representation of the matrix Z is shown in the equation 1.

$$Z = \begin{array}{c c} & \begin{array}{cccc} Z_1 & Z_2 & \cdots & Z_n \end{array} \\ \begin{array}{c} t_1 \\ t_2 \\ \vdots \\ t_m \end{array} & \begin{pmatrix} z_{11} & z_{12} & \cdots & Z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{m1} & z_{m2} & \cdots & z_{mn} \end{pmatrix} \end{array} \quad (1)$$

To distinguish abnormal values from normal values, we are using Back Propagation Neural Network (BPNN) algorithm, once abnormal values for physiological parameters are detected, Ensemble LinReg is used to predict the current value for that particular record using the other co-related parameters, and when the difference between the actual and predicted parameter is greater than the pre-specified threshold, then anomalous data is differentiated from genuine data using a correlation analysis.

In the remaining section, we are going to briefly discuss about the BPNN and Ensemble LinReg method, and how they are used in our proposed model. Please refer [18], for further details about aforementioned models.

A. BPNN

BPNN stands for back propogation neural network. In this neural network model, information flows forward from input layer to the subsequent layers, whereas the error is propogated backwards, hence, it is termed as feed-forward back propogation neural network. It consists of one or more hidden layer in the middle of one input and one output layer. For our model of BPNN, we are using one hidden layer in between of input and output layer. Architecture of the BPNN model is shown in figure 2.

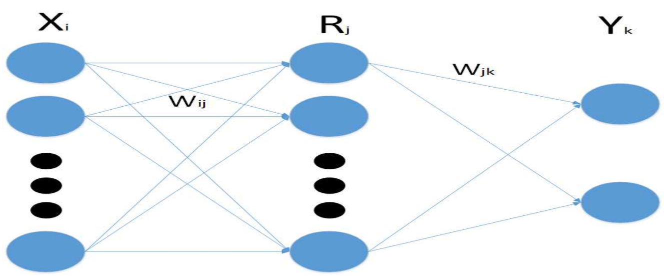

Fig. 2. Three layered Feed Forward Back Propagation Neural Network

BPNN model is a supervised learning algorithm, in which neural network output, i.e. output obtained from the output layer is compared to actual output by calculating the least mean square error and following that the randomly assigned weights are modified with each iteration to reduce the error as minimal as possible.

Here, Z_i represents inputs fed to input layer, for our dataset, physiological parameters are fed to input layer. R_j represents the output obtained from the hidden layer, Y_k represents the output obtained from the neural network(output layer), i.e. physiological parameters are classified into normal and abnormal values. W_{ij} and W_{jk} represents the connection weights.

We are using Gradient Decent Method (GDM) to reduce the mean squared error between output obtained from neural network and actual output. With the combination of TRAINSCG for training, LEARNGDM for learning to calculate output R_j for every i^{th} node in the hidden layer and LOGSIG for transfer functions to calculate output Y_k, the BPNN algorithm is successful in the classification for abnormal and normal sensed data.

TRAINSCG is function available in MATLAB library, which is a training function used to revise weight and bias values as per the scaled conjugate gradient method.

LEARNGDM, available in MATLAB Library, which is used to find out the change in weight dW for any particular neuron from the model's input layer z_{ij}, momentum constant *m*, bias (or weight) *W*, learning rate l_r, as per the gradient descent method.

$$dW = m \times dW_{prev} + (1 - m) \times l_r \times gW \tag{2}$$

dW_{prev} is the previous change in weight, which is stored and read from the learning state l_s .

LOGSIG is a transfer function available in MATLAB library, which evaluates a layer‘s output obtained from its net input.

$$logsig(x) = \frac{1}{1 + exp^{-x}} \tag{3}$$

B. Bagging on Linear Regression

Linear regression is a machine learning technique [19] [18] which is used to predict the current value for dependent attribute z_{ij} in instance *i* using other correlated attributes y_{ij}, which are referred as regressors, which can be statistically defined as:

$$z_{ij} = C_0 + C_1 \times y_{i1} + C_2 \times y_{i2} + C_3 \times y_{i3} + ... + C_n \times y_{in} \tag{4}$$

where C_k $\forall$ k = 1,2,3,..,n are the weights or generally referred as coefficients of regressors. At the training phase, the weights are calculated using covariance of Z_i and Y_i divided by the variance of Y_k.

$$C_k = \frac{Cov(Zk, Y_k)}{Var(Y_k)} = \frac{\sum(z_{ik} - \bar{Z_k})(y_{ik} - \bar{Y_k})}{\sum(y_{ik} - \bar{Y_k})} \tag{5}$$

We are using Ensemble LinReg to predict the current value for the dependable monitored attributes z_{ik} by using the correlated independent parameters of the same instance $y_{ij|j \neq k}$. Later, check whether the predicted value falls within the pre-specified small margin of error.

Bagging, also referred as Bootstrap Aggregating is statistical estimation method. It is a ensemble machine learning based technique which can be applied on any classification or regression model. In Bagging, numerous random samples of the training data are drawn with substitution and used to train various distinctive base models, in our experiment on Linear Regression, which results to give more robust prediction by taking average values from all the different base models. Bagging on Linear Regression results in higher accuracy, efficiency and robust predictions than the original LinReg model.

The proposed algorithm 1 consist of two step machine learning process, in first step every record is classified as normal and abnormal. BPNN is used to classify the abnormality of the record, and a physiological parameter is considered abnormal if it falls outside the pre-specified normal range for that parameter, such as, ABPmeans $\in [70, 120]$, PULSE $\in [60, 120]$, RESP $\in [15, 30]$, HR $\in [60, 120]$, $SpO_2 \in [90, 100]$. HR and PULSE are similar attributes obtained from two different sensors, like, PULSE is sensed from Pulse Oximeter and HR is evaluated from the number of interbeat intervals (R-R) in ECG signals.

```
foreach received record R_i during T do
    Classify R_i using BPNN;
    if Class (R_i) = 'ABNORMAL' then
        foreach y_ik do
            z_ik = sum_{j=1,j≠k}^{n} C_j y_ij ;
            counter += (|y_ik − z_ik| ≥ 0.1 × z_ik)?1 : 0;
        end
        if counter ≥ k then
            Alert healthcare professional;
        end
    end
end
```

Algorithm 1: Anomaly Detection Algorithm

For each record, classified as abnormal, Ensemble LinReg predicts the current value for the dependable attribute of that instance using the other correlated attributes. To check whether the sensed value is genuine or anomalous, we are using the below mentioned equation 6.

$$e_i = |y_{ik} - z_{ik}| \geq 0.1 \times z_{ik} \tag{6}$$

Our proposed approach is mainly divided into three different phases, first is offline training, after that online testing and then finally detection. For the initial offline training phase, BPNN model is trained to classify the physiological parameters, and In online testing phase, inputs are classified as normal or abnormal by BPNN model, if they lie outside the normal range specified for each parameter. In our analysis, BPNN is proven to be most efficient as compared to already existing J48 [6] and SVM model [17].

Later, for abnormal classified attributes, we are using Ensemble LinReg to predict the current value using the correlated attributes. For example, if an instance of HR values (y_{ik}) are classified as abnormal, we are recursively assuming for that particular instance HR value is missing, hence, using Bagging on LinReg, we can predict the current HR value (z_{ik}) using other correlated parameters as PULSE, RESP rate, SPo2 and ABPmeans $(z_{ij|j\neq k})$.

$$\hat{HR_i} = C_0 + C_1 \times Pulse_i + C_2 \times RESP_i + C_3 \times SPo2_i + ... \tag{7}$$

If the Euclidean Distance (error) between actual HR value HR_i and predicted value $\hat{HR_i}$ is greater than the pre-specified threshold (10% of the predicted value), that instance of the record is considered as anomalous, and replaced by the predicted value, found using Bagging on LinReg. However, if k readings are greater than pre-specified threshold, we raise the alarm for the Healthcare Professional, to alert him/her or about the deteriorating conditions of patient, for example, heavy fluctuation in HR values or ABPmeans value are serious health condition, and require immediate attention and proper diagnosis. Instead of a fixed amount of readings, we are considering variable value for k, as the fluctuations of health parameters can also be caused due to medication from healthcare professionals. In our experiment, we are assuming the value of k as 3.

IV. RESULTS

In this section, we will show the results obtain after testing our proposed model for Anomaly Detection in WBAN sensors. Afterwards, we will analyse the performance of our model using various parameters like Mean Absolute Error Rate(MAE), Root Mean Squared Error(RMSE), TPR, FPR. In order to validate robustness of our model, we are using Real Patient's Data obtain from Physionet Database [20], which comprises of 85983 records. Each record contains 5 parameters HR, ABPmeans, PULSE, RESP, SpO2. The variations in values of HR, PULSE, RESP rate are shown in figure 3 and variations of values of SPo2 and ABPmeans are shown in figure 4.

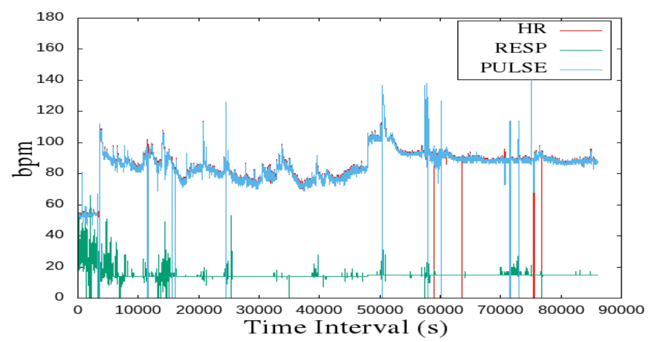

Fig. 3. Variation of HR, PULSE & RESP

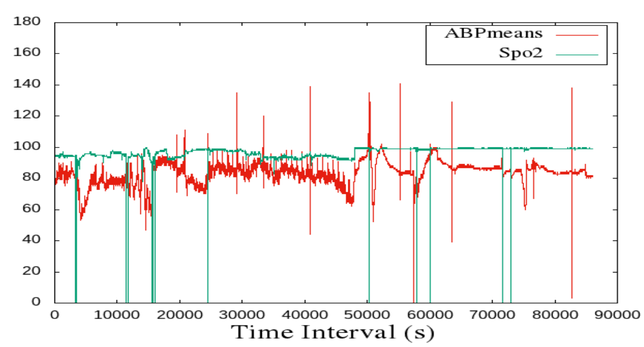

Fig. 4. Variation of SPo2 & ABPmeans

In our experiment, we are assuming, HR values sensed by the ECG sensor, might give anomalous data, hence, we will be focussing on the predicted values of HR only. With the help of WEKA Tool [21], Figure 5 shows the actual HR, predicted HR and error (Euclidean Distance) between the actual HR value and predicted HR value computed using Linear Regression. To validate our proposed approach, we are comparing the results with other classification models as well: Linear Regression 6, REPTree(Reduced Error Pruning Tree) 7, Additive Regression 8, Multilayer Perceptron 9.

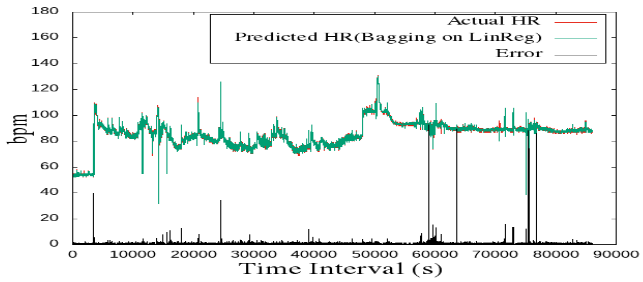

Fig. 5. Bagging on Linear Regression Model

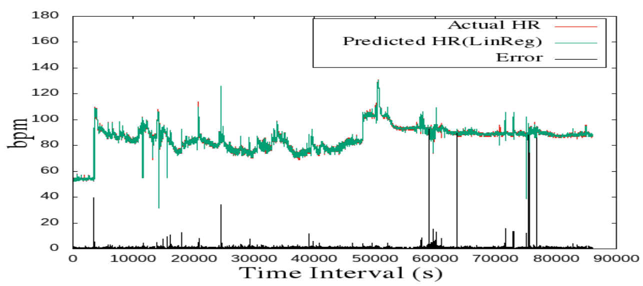

Fig. 6. Linear Regression Model

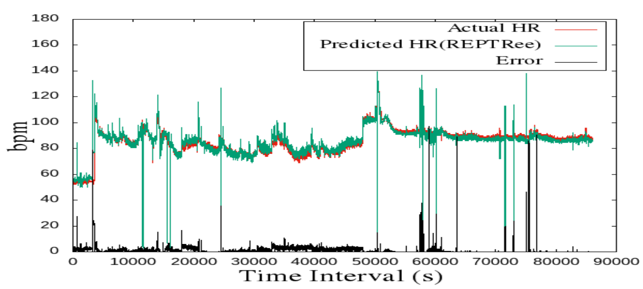

Fig. 7. Reduced Error Pruning Tree Model

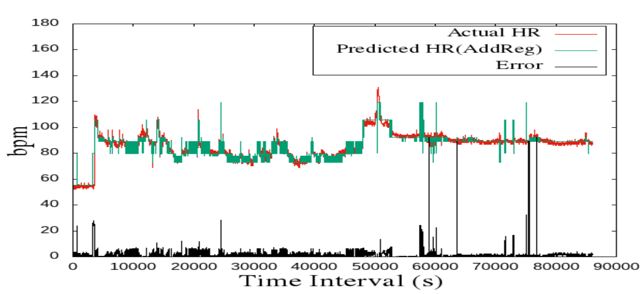

Fig. 8. Additive Regression Model

In Table I, we show the comparison of various performance metrics(Accuracy, Mean Absolute Error(MAE), Root Mean Square(MSE)) for the classification models used, as we can

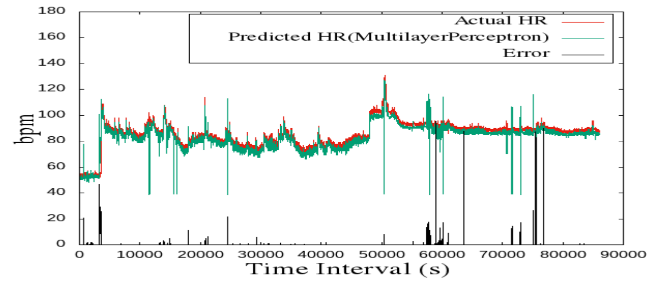

Fig. 9. Multilayer Perceptron Model

TABLE I
PERFORMANCE ANALYSIS

Models \ Performance Metrics	Accuracy	MAE	RMS
Bagging on Linear Regression	0.985	0.570	1.770
Linear Regression	0.971	0.681	1.883
Additive Regression	0.940	1.98	3.535
REPTree	0.931	2.439	3.742
Multilayer Perceptron	0.957	2.908	3.99

see the Bagging on LinReg shows the best result among all. Bagging on LinReg had the lowest MAE & RMS as well as maximum accuracy out of all other classifications models, which is why we have used this model for prediction in our approach.

Figure 10, shows the alarms raised by our approach of BPNN with Bagging on LinReg. Alarms are raised because of heavy changes in atleast *k* attributes. On comparing our results with the existing approaches in which BPNN with Bagging on LinReg is replaced by J48 with LinReg and SVM with LinReg. We can see almost all the three figures 10, 11 & 12 shows similar results, whereas J48 with LinReg model shows some extra raised alarms than SVM with LinReg(in the beginning before 10000 seconds), and SVM with LinReg shows some extra alarms on comparison to our model(in the beginning before 10000 seconds, before 60000 seconds and after 80000 seconds). This proves that, our proposed approach is able to minimize the FPR for the Anomalies for WBAN sensor.

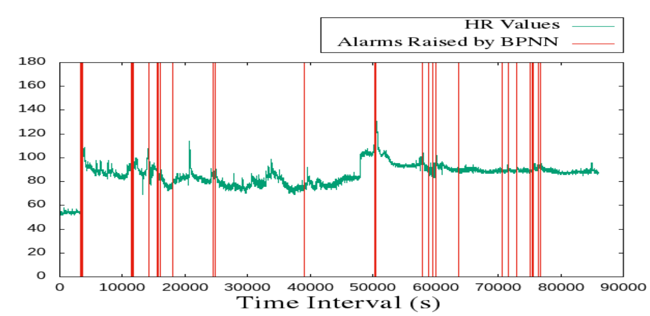

Fig. 10. Alarms raised by BPNN Model

Alongside, the above performance metrics, two other performance metrics TPR and FPR are used to validate our approach, and also for comparing our approach with existing approach. Value for the TPR and FPR can be calculated by the following equations 8 and 9 respectively.

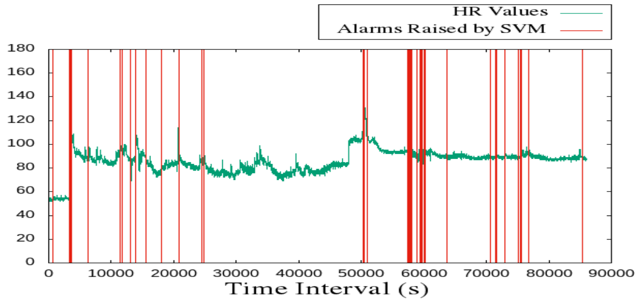

Fig. 11. Alarms raised by SVM Model

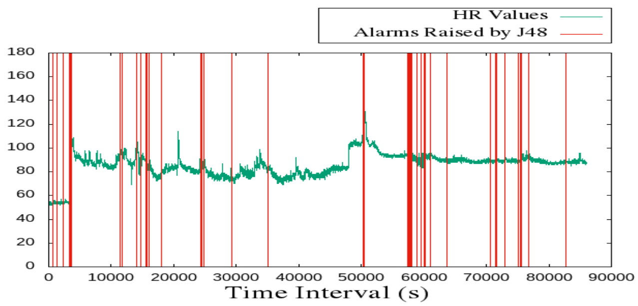

Fig. 12. Alarms raised by J48 Model

$$TPR = \frac{TP}{TP + FN} \quad (8)$$

$$FPR = \frac{FP}{FP + TN} \quad (9)$$

BPNN, SVM and J48 classification models are the three most accurate machine learning algorithms used for classification of abnormalities and their combination with LinReg is best algorithm for Anomaly Detection in WBAN sensor with True Positive Rate of 100% and False positive rare 4.2%, 6.5% and 7.4% respectively. The above results validates our claim of our approach proves to be robust and highly efficient in comparison to the existing approaches for detecting anomalies in WBAN sensor.

V. Conclusion

In this paper, we have integrated ANN with Ensemble LinReg to propose a detection algorithm for anomalies in WBAN Sensors. The proposed approach is able to get us result with high TPR and least FPR on comparing with the pre-existing approaches. We have used the real patient‘s data with over a large time interval to train, test and predict and finally detect whether there exist a anomaly in the data or not. Our proposed algorithm is able to reduce the false alarm rate with high anomaly detection accuracy.

References

[1] P. Kumar, S.-G. Lee, and H.-J. Lee, “E-sap: efficient-strong authentication protocol for healthcare applications using wireless medical sensor networks,” *Sensors*, vol. 12, no. 2, pp. 1625–1647, 2012.

[2] C. Otto, A. Milenkovic, C. Sanders, and E. Jovanov, “System architecture of a wireless body area sensor network for ubiquitous health monitoring,” *Journal of mobile multimedia*, vol. 1, no. 4, pp. 307–326, 2006.

[3] J. Ko, J. H. Lim, Y. Chen, R. Musvaloiu-E, A. Terzis, G. M. Masson, T. Gao, W. Destler, L. Selavo, and R. P. Dutton, “Medisn: Medical emergency detection in sensor networks,” *ACM Transactions on Embedded Computing Systems (TECS)*, vol. 10, no. 1, p. 11, 2010.

[4] H. Wang, H. Fang, L. Xing, and M. Chen, “An integrated biometric-based security framework using wavelet-domain hmm in wireless body area networks (wban),” in *Communications (ICC), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–5.

[5] Y. Zhang, N. A. Hamm, N. Meratnia, A. Stein, M. Van De Voort, and P. J. Havinga, “Statistics-based outlier detection for wireless sensor networks,” *International Journal of Geographical Information Science*, vol. 26, no. 8, pp. 1373–1392, 2012.

[6] O. Salem, A. Guerassimov, A. Mehaoua, A. Marcus, and B. Furht, “Sensor fault and patient anomaly detection and classification in medical wireless sensor networks,” in *IEEE International Conference on Communications (ICC), 2013*. Citeseer, 2013, pp. 4373–4378.

[7] D. Malan, T. Fulford-Jones, M. Welsh, and S. Moulton, “Codeblue: An ad hoc sensor network infrastructure for emergency medical care,” in *International workshop on wearable and implantable body sensor networks*, vol. 5. Boston, MA;, 2004.

[8] K. Montgomery, C. Mundt, G. Thonier, A. Tellier, U. Udoh, V. Barker, R. Ricks, L. Giovangrandi, P. Davies, Y. Cagle *et al.*, “Lifeguard-a personal physiological monitor for extreme environments,” in *Engineering in Medicine and Biology Society, 2004. IEMBS’04. 26th Annual International Conference of the IEEE*, vol. 1. IEEE, 2004, pp. 2192–2195.

[9] A. Wood, G. Virone, T. Doan, Q. Cao, L. Selavo, Y. Wu, L. Fang, Z. He, S. Lin, and J. Stankovic, “Alarm-net: Wireless sensor networks for assisted-living and residential monitoring,” *University of Virginia Computer Science Department Technical Report*, vol. 2, p. 17, 2006.

[10] K. F. Navarro, E. Lawrence, and B. Lim, “Medical motecare: A distributed personal healthcare monitoring system,” in *International Conference on eHealth, Telemedicine, and Social Medicine*. IEEE, 2009, pp. 25–30.

[11] J. P. S. Cunha, B. Cunha, A. S. Pereira, W. Xavier, N. Ferreira, and L. Meireles, “Vital-jacket®: A wearable wireless vital signs monitor for patients’ mobility in cardiology and sports,” in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010*. IEEE, 2010, pp. 1–2.

[12] H. Alemdar and C. Ersoy, “Wireless sensor networks for healthcare: A survey,” *Computer networks*, vol. 54, no. 15, pp. 2688–2710, 2010.

[13] K. Grgić, D. Žagar, and V. Križanović, “Medical applications of wireless sensor networks-current status and future directions.” *Medicinski Glasnik*, vol. 9, no. 1, 2012.

[14] Y.-Y. Zhang, H.-C. Chao, M. Chen, L. Shu, C.-H. Park, and M.-S. Park, “Outlier detection and countermeasure for hierarchical wireless sensor networks,” *IET Information Security*, vol. 4, no. 4, pp. 361–373, 2010.

[15] S. Rajasegarar, C. Leckie, J. C. Bezdek, and M. Palaniswami, “Centered hyperspherical and hyperellipsoidal one-class support vector machines for anomaly detection in sensor networks,” *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 3, pp. 518–533, 2010.

[16] S. Xu, C. Hu, L. Wang, and G. Zhang, “Support vector machines based on k nearest neighbor algorithm for outlier detection in wsns,” in *Wireless Communications, Networking and Mobile Computing (WiCOM), 2012 8th International Conference on*. IEEE, 2012, pp. 1–4.

[17] O. Salem, A. Guerassimov, A. Mehaoua, A. Marcus, and B. Furht, “Anomaly detection in medical wireless sensor networks using svm and linear regression models,” *International Journal of E-Health and Medical Communications (IJEHMC)*, vol. 5, no. 1, pp. 20–45, 2014.

[18] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.

[19] C. M. Bishop, “Pattern recognition and machine learning (information science and statistics) springer-verlag new york,” *Inc. Secaucus, NJ, USA*, 2006.

[20] “”Physionet. http://www.physionet.org/cgi-bin/atm/atm”.”

[21] “”Weka Data Mining Tool. http://www.cs.waikato.ac.nz/ml/weka/”.”