International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 09 | September - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

INVESTIGATION ANALYSIS ON E-HEALTHCARE MONITORING SYSTEM WITH IOT AND BIG DATA BY TSLR-CHSC MODEL

V. Deepa [1]

Ph. D (Research Scholar)[1]

PG and Research Department of Computer Science,

Tirppur Kumaran College for Women, Tirppur, Tamil Nadu, India

Email id: deepavmsc@gmail.com

Dr. K. Rajeswari [2]

Associate Professor [2]

PG and Research Department of Computer Science,

Tirppur Kumaran College for Women, Tirppur, Tamil Nadu, India

Email id: rajeswarik1224@gmail.com

Abstract: - Internet of Things (IoT) technology helped the development of E-Health care Monitoring System from direct visit to virtual Monitoring (Telemedicine). Smart health care system in IoT environment monitored the basic health signs such as heartbeat rate, body temperature and blood pressure in real-time applications. IoT and big data is a prominent challenge in smart healthcare system. In this health monitoring system big data is employed to analyze the large volume of data and to determine the Normal and Abnormal patient condition. Numerous issues like accuracy, time and error have yet to be conveyed to generate a ductile system for health care monitoring. To address these issues, I proposed a method called Theil Sen Linear Regression and Canopy Hopkins Statistics Clustering (TSLR-CHSC) for IoT based Health monitoring system is proposed. This method splits into three sections such as Data collection, Feature Selection and Clustering. First, Cardiovascular disease dataset is acquired from sensors are collected. Second, appropriate features can be selected by using Theil Sen regression feature. Third, clustering is performed based on the cluster tendency by using canopy algorithms. Through this way I deployed an efficient E-Health Monitoring System with minimum time consumption. For evaluation, a cardiovascular disease dataset is obtained from various medical sensor devices are analyzed to identify the disease severity.

Keyword: *IoT, Big Data, Theil Sen Linear Regression, Canopy Hopkins Statistic Clustering, Health Care Monitoring.*

I. INTRODUCTION

In recent years, the initiation of Internet of Things (IoT) helps the progress of healthcare monitoring from face- to-face consulting to telemedicine. IoT is an environment where each node is connected to other node that relates to the network to transfer data for taking decision. An IoT based healthcare monitoring system check the vital signs of both biological and behavioral changes of the patients And a PATH2Iot Framework was introduced to decompose complex IoT applications into smaller (micro) operations.

To analyze the data a machine learning technique EEPSOC (Energy Efficient Particle Swarm Optimization) based clustering method is implemented. In this method, the sensed data acquired from the dataset were grouped based on its Cluster Head (CH) value. Now the appropriate CH Data is transmitted over to the cloud server. Here ANN (Artificial Neural Network) based classification model is designed to diagnose disease severity from the patient dataset. This designed mechanism consisting of three phases such as Emergency detection, Adapting sensing frequency and real time patient situation prediction. If the sensor nodes failed to avoid repeated collision, then the designed mechanism could not adjust the sensing frequency.

II. LITERATURE SURVEY

Healthcare monitoring is a crucial parameter with respect to accuracy, time and error involved in prediction of several diseases like, diabetes and cardiovascular disease. The objective is to propose a health care monitoring system which will improve the QoS in terms of clustering accuracy, clustering time and error. For IoT with sensors collecting patient's data, large amount of data is said to be created by sensors while extracting healthcare data. However, a machine learning technique can be utilized to learn the patient's data and process accordingly, however, suffer from serious issues, henceforth. affecting time, accuracy and error involved in disease prediction. In such circumstances to monitor the cardiovascular disease with cluster tendency aspect, a cluster threshold is utilized to map the cluster tendency with the actual cluster formation. An IoT health care monitor system has been proposed based on machine learning for cardiovascular disease prediction in case of large involvement of data (i.e., Big Data).The patient's healthcare monitor based cardiovascular disease prediction model is proposed to advice the human being affected with disease regarding further treatment, to have control on certain aspects concerning disease and son. The healthcare monitoring method shows better accuracy than the existing methods. A Theil–Sen Estimated Linear Regression Feature Selection model is first designed to select the relevant features or attributes for cardiovascular disease monitoring via multiple linear regressions and Theil–Sen Estimator function.

A Canopy Hopkins Statistic Clustering algorithm for healthcare monitoring is designed with the selected feature for detecting either the presence or absence of disease by employing appropriate learning parameters. The prediction accuracy, prediction time and error for different numbers of samples or patients while using the proposed method are determined and experimental analysis is evaluated. a new method of wearable sensor device for acquiring real time athlete data using IoT for monitoring electrocardiogram (ECG) patterns in addition to the body acceleration using smart phone was proposed. Also, with the obtained data classification using Radial-basis Function Network and Levenberg-Marquardt with Probabilistic Neural Network was performed for health monitoring. Some of the issues and gaps in IoT for healthcare monitoring were analyzed.

III. E-HEALTHMONITORING SYSTEM WITH IOT AND BIGDATA

Internet is used in different fields like education, finance, business, healthcare, industries and social networking, shopping, E-commerce. A modernized healthcare system present improved healthcare services to the people at anytime and anywhere. Healthcare Monitoring System (HMS) is employed to arrange day-to-day process with the system activities such as clinical, administration and commercial perspectives. HMS is constructed on subsystems to obtain tolerable and clinical management. The information is gathered from different sources such as patient health data, hospital administrative data and external data.

IV. THEIL SEN LINEAR REGRESSION AND CANOPY HOPKINS STATISTIC CLUSTERING FOR IOT -BASED HEALTHCARE MONITORING

Internet of Things and Big data, information can be interchanged swiftly and in a more secure manner. The Internet of things (IoT) with Big Data provides numerous opportunities in the field of e-health care monitoring. Using IoT with Big Data in this process can notably enhance the patient monitoring. Hence, it becomes paramount to provide a useful method in the medical industry to monitor the patients' status employing associated sensors. In this section we plan to develop a method called, Theil Sen Linear Regression and Canopy Hopkins Statistic Clustering (TSLR-CHSC), for IoT-based healthcare monitoring with higher accuracy and lesser time consumption. The proposed TSLR-CHSC method is split into three sections.

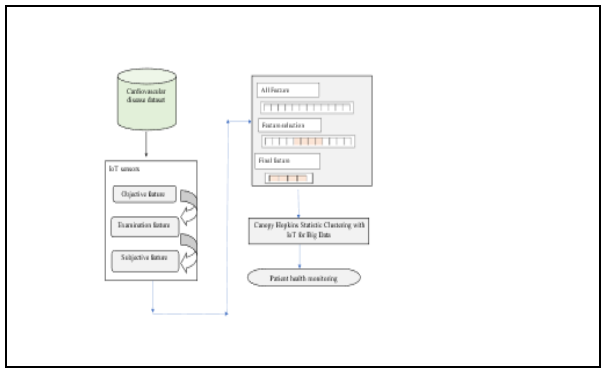

Fig 1: **The structure Theil Sen Linear Regression and Canopy Hopkins Statistic Clustering (TSLR-CHSC) method.**

This TSLR-CHSC method split into three sections. In First section, IoT patient data is collected from the cardiovascular disease dataset. In Second section, relevant features or attributes can be selected by using Theil Sen Estimated Linear Regression Feature. In third section, Clustering is carried out to disease diagnose based on the error rate can be obtained.

A. Theil Sen Linear Regression Feature Selection Model

In this section, the objective, subjective and examination feature can be selected based on IoT patient data that is collected from the database is selected with minimum time and maximum accuracy. Bayesian Linear function is used to find the median of slope and median of intercept with their relevant features.

Input: Dataset '', Features ''
Output: Computationally efficient and accurate feature selection ''
Step 1: Begin
Step 2: For each Dataset '' with Features ''
Step 3: Formulate multiple linear regressions as in equation (1)
Step 4: Estimate slope of the corresponding feature sample points '' in the plane as in equation (2)

Step 5: Evaluate sum of squares of residuals as in equation (3)
Step 6: Evaluate least square estimators as in equation (4)
Step 7: Estimate intercept of the corresponding feature sample points '' in the plane as in equation (5)
Step 8: Evaluate intercept median of the least square estimators as in equation (6)
Step 9: Estimate Bayesian linear regression with respect to the slope and intercept as in equations (7) and (8)
Step 10: Return relevant features '' as in equation (9)
Step 11: End for
Step 12: End

.

Algorithm 1 Theil Sen Linear Regression Feature Selection

Here multiple linear regression is formulated for distinct features present in the dataset. Now, the median of slope (i.e., to analyses the cardiovascular disease) to the sample points (i.e., values of the features) in the plane (i.e., for overall dataset) is estimated. Finally, the optimal feature is obtained by employing Linear function.

B. Canopy Hopkins Statistics Clustering for Healthcare Monitoring

In this section, Clustering process is carried out to diagnose the disease in an effective manner. Here Canopy Clustering is an unsupervised Learning. This clustering process initializes two clusters, one for Normal Patient and the other for Abnormal patients. Hopkins statistics analysis is carried out to determine the cluster tendency to what degree clusters exist in data to be clustered. Here clustering tendency has to be evaluated to analyze whether the feature selected form cardiovascular disease dataset possess meaningful clusters or not. To obtain accurate clustering the conditional checking is made between the sampled distance and threshold and based on value the cluster is assigned as Normal. On the other hand, conditional checking is made between the random distance and the threshold, and this cluster assigned as Abnormal condition. This process is repeated until there is no more datapoints are left in the dataset.

V. EXPERIMENTAL EVALUATION

Experimental Evaluation were carried out on certain parameters like clustering accuracy, clustering time and error rate for sample patient data collected at different time intervals. Several performance metrics were utilized to estimate the efficiency of the TSLR-CHSC method. Clustering Accuracy represents the overall clustering potentiality of the proposed machine learning method.

Number of patient data	TSLR-CHSC	MACHINE LEARNING BASED HEALTHCARE	PATHlist
7000	97.21	96.5	95.5
14000	95.15	94.35	93.15
21000	94	92	90.55
28000	93.25	90.35	89.25
35000	93	88.15	87.35
42000	92.85	87.35	85
49000	92.55	86.85	84.15
56000	92	85.35	83.15
63000	90	85.15	82
70000	88.35	85	82

Table 1 performance evalution of clustering accuracy

IMPACT ON ACCURACY: Table 1 depict the Clustering Accuracy is described as the ratio of number of patient data that are correctly classified the patient health condition to the total number of patient data considered as input.

Table 1 shows the result obtained by TSLR-CHSC Model is as follows.

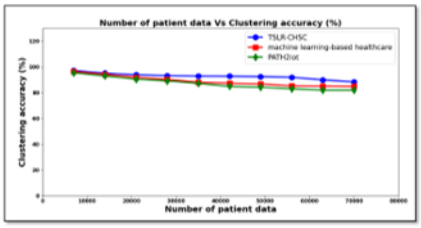

Fig 2 Graphical representation of clusterig accuracy

Then the accuracy value is calculated. when the accuracy value is higher, the method is said to be more efficient. Accuracy is maximized to 82%. while using TSLR-CHSC model.

IMPACT ON TIME COMPLEXITY: Time complexity is measured as an amount of time consumed by the algorithm to classify Normal and Abnormal patients. Time complexity is measured in milliseconds. Time complexity is minimized to 19% by using TSLR-CHSC model.

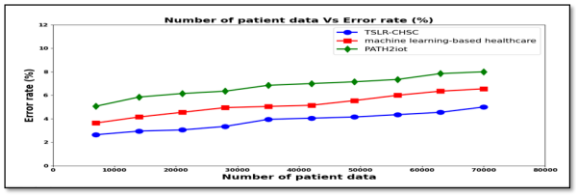

Fig 3: Graphical representation of Error rate

IMPACT ON ERROR RATE: Fig 3 depict the Error rate is measured based on the ratio of several patient data that are incorrectly classified into the class to the total number of patient data. The error rate is minimized to 8% by using TSLR-CHSC model.

CONCLUSION

In this paper, machine learning based Theil Sen Linear Regression and Canopy Hopkins Statistic Clustering (TSLR-CHSC) method for IoT with Big Data is proposed to check vital features concerning cardiovascular disease and monitor the changes of persons via smart personnel care technologies. According to the theoretical model of the proposed method, two sections have been considered. These sections include obtaining the relevant feature or attribute for healthcare monitoring concerning cardiovascular disease employing Theil–Sen Estimator function and clustering via Canopy Hopkins Statistic Clustering algorithm. The proposed method was evaluated with different clustering methods. The applied clusters included machine learning and linear regression models. The experimental results revealed that the clustering algorithms using machine learning, called, TSLR-CHSC method performed well in terms of the clustering accuracy, clustering time and error rate. TSLR-CHSC method reached the highest performance for healthcare monitoring in our scenario with 97.21% accuracy, 1085ms time and error rate of 2.64%. High accuracy of TSLR-CHSC method in comparison with other applied clustering methods is a significant difference that makes it applicable in real-time healthcare monitoring.

REFERENCES

[1] Alireza Souri, Marwan Yassin Ghafour, Aram Mahmood Ahmed, Fatemeh Safara, Ali Yamini and Mahdi Hoseyninezhad, “A new machine learning-based healthcare monitoring model for student’s condition diagnosis in Internet of Things environment”, Soft Computing, Springer, Volume 24, 2020, Pages 17111–17121 [machine learning-based healthcare]

[2] Devki Nandan Jha, Peter Michalak, Zhenyu Wen, Rajiv Ranjan, and Paul Watson “Multi-objective Deployment of Data Analysis Operations in Heterogeneous IoT Infrastructure”, IEEE Transactions on Industrial Informatics, Volume 16, Issue 11, November 2020, Pages 7014 – 7024 [PATH2iot]

[3] R. Bharathi, T. Abirami, S. Dhanasekaran, Deepak Gupta, Ashish Khanna, Mohamed Elhoseny K. Shankar, “Energy Efficient Clustering with Disease Diagnosis Model for IoT based Sustainable Healthcare Systems”, Sustainable Computing: Informatics and Systems, Volume 28, December 2020, Pages 1-28

[4] Hassan Harb, Ali Mansour, Abbass Nasser, Eduardo Motta Cruz and Isabel de la Torre Diez, “Sensor-Based Data Analytics for Patient Monitoring in Connected Healthcare Applications”, IEEE Sensors Journal, Volume 21, Issue 2, January 2021, Pages 1-10

[5] Wei Li, Yuanbo Chai, Fazlullah Khan, Syed Rooh Ullah Jan, Sahil Verma, Varun G. Menon, Kavita, Xingwang Li, “A Comprehensive Survey on Machine Learning-Based Big Data Analytics for IoT-Enabled Smart Healthcare System”, Mobile Networks and Applications, Springer, Jan 2021

[6] Xingdong Wu, Chao Liu, Lijun Wang, Muhammad Bilal, “Internet of things-enabled real-time health monitoring system using deep learning”, Neural Computing and Applications, Springer, Aug 2021

[7] Somayeh Iranpak, Asadollah Shahbahrami, Hassan Shakeri, “Remote patient monitoring and classifying using the internet of things platform combined with cloud computing”, Journal of Big Data, Aug 2021

[8] Heng Yu, Zhiqing Zhour, “Optimization of IoT-Based Artificial Intelligence Assisted Telemedicine Health Analysis System”, IEEE Access, Jun 2021