REMOTE HEALTHCARE MONITORING METHODOLOGY FOR DIABETES PATIENTS USING LOGISTIC REGRESSION OVER DECISION TREE FOR IMPROVING PRECISION

M Nishanth Naren[1], V. Karthick[2]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: The goal of this research is to self monitor in remote healthcare and help in diabetic prediction for diabetes patients.

Materials and Methods: There are two groups in this study. The first group built a Logistic Regression technique, whereas the second group constructed a Decision Tree with 104 samples. The sample size for Logistic Regression is 52, with a Decision Tree (N = 52) and G power (value = 0.8) sampling strategy.

Result: The accuracy of the Decision Tree Method has been enhanced to 86%, while the accuracy of the Logistic Regression algorithm has been shown as 79%. The mean accuracy detection is ±2SD and the significance value is 0.000 ($p<0.05$) which shows the hypothesis is correct and it is carried out using an independent sample T test.

conclusion: The Decision Tree Approach's final result (86%) was determined to be much more accurate than the Logistic Regression algorithm (79%).

Keywords: Logistic Regression, Decision Tree, Novel Healthcare, Self monitoring, Diabetic prediction.

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu,India. Pincode: 602105.

[2*]Department of Computer Science and Engineering, Saveetha School Of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India. Pincode: 601205.

1. Introduction

Healthcare is a rapidly evolving area in terms of technology and services. Remote patient self monitoring is a recent development in this field that has several advantages in an aging world population with increasing health issues and healthcare (Costache et al. 2021). With relatively easy applications to monitor patients inside hospital rooms, technology has progressed to the point where patients can go about their everyday lives at home while still being observed using modern communication and sensor technologies (Costache et al. 2021; Brooks et al. 2021). Sensors for EKG readings, heart rate, respiration rate, blood pressure, temperature, blood glucose levels, diabetic prediction and brain system activity are now accessible (Niu et al. 2021). The scope of remote Novel healthcare includes everything from Self monitoring chronically ill patients healthcare, elders healthcare, premature children healthcare,and accident victims healthcare (Niu et al. 2021; Patel et al. 2021).

More than 460 articles were published in this diabetic prediction system for the past five years from various sources such as Google Scholar, IEEE Xplore and Springer.This new technology can keep track of patients based on their sickness or their circumstances (Tao et al. 2021; Thenkabail 2015). As the globe evolves toward remote Self monitoring and real-time and fast diagnosis of ailments, Novel healthcare is becoming an emerging research subject.There are several types of remote novel healthcare,all of which include using technology to monitor patients outside of hospital settings (Walton 2010). Our team has extensive knowledge and research experience that has translated into high quality publications (Pandiyan et al. 2022; Yaashikaa, Devi, and Kumar 2022; Venu et al. 2022; P. G. Kumar et al. 2022; Nagaraju et al. 2022; Karpagam et al. 2022; Baraneedharan et al. 2022; Whangchai et al. 2022; Nagarajan et al. 2022; Deena et al. 2022)

The advantages of remote Self monitoring of patients are: early and real-time detection of illnesses; ability to continuously monitor patients and prevention of worsening of illnesses and untimely deaths; cost reduction in hospitalizations; and reducing the number of hospitalizations. Research gap of this paper is that while obtaining more accurate readings fast diagnosis of ailments faces more challenges. While permitting usual daily activities, patients efficiency in novel healthcare services by utilizing communication technology and emergency medical care; service for patients with mobility issues; emergency care for traffic accidents and other injuries; and usage of non-invasive medical interventions (A. Kumar and Mozar 2020). The aim of the novel innovative healthcare framework is to provide diabetes patients to notify instructors to improve speed and accuracy.

2. Materials and Methods

This study was conducted at the Saveetha Institute of Medical and Technical Sciences' Department of Artificial Intelligence Laboratory in the Saveetha School of Engineering. (Tao et al. 2021). The Logistic Regression method (79%) and Decision Tree (86% are the two groups in this study 86%. As a whole, there are two groups in this project. The existing system is referred to as Group 1, and the new system is referred to as Group 2. There are a total of 104 samples, out of which 52 are samples for the first group, and the remaining 52 are samples for the second group (Ebben et al. 2022).

Same dataset has been utilized for both the techniques. All empty values in the dataset are examined and validated. In order to utilize and install all of the necessary software, a minimum of 4GB of RAM as well as an i3 processor are required to run all of the applications simultaneously (Sullivan 2018). At least 50GB hard disc space and an internet connection to download and install all of the necessary tools and development environment to operate this Novel Effective framework is needed. The application is written in the Python programming language . The IDLE is used to start and execute the application, which is written in Python 3.7.

Logistic Regression Algorithm

Logistic regression, a supervised learning technique, is used to predict a dependent categorical target variable. To put it another way, if they have a vast volume of data to categorize, provision regression might be able to help using Novel Healthcare. If they were given a dog and an associate's degree orange and asked to classify them both as animals, then they might probably say the dog is an associate's degree animal and the orange isn't. The goal is to accurately classify an associate's degree animal, and have to rely on our own knowledge to do so. Logistic regression is one of the most well-known machine learning algorithms in the supervised learning technique in Self monitoring. Exploitation of categorical dependent variable's output is predicted via provision regression. As a result, the outcome should be categorical or distinct. It's often positive or negative, 0 or 1, true or false, and so on. But, instead of precise values like zero and one, it gives probabilistic values in the center. Provision regression and rectilinear regression are very similar, except for how these are used in this

methodology. Provision regression questions are solved using rectilinear regression, while classification challenges are solved using provisional regression.

Logistic Regression steps:
Gathering Data is the first step.

Step 2: analyze the data set for training.Using Logistic Regression
Step 3: Test Predicts the Outcomes
Step 4: creation of the confusion matrix and check
Step 5: visualize the accuracy of the algorithm
Step 6: output is displayed.

Decision Tree Algorithm
Decision Trees are a type of supervised machine learning in which the data is continually split according to a parameter (you explain what the input is and what the related output is in the training data)in Self monitoring. Two entities, decision nodes and leaves, can be used to explain the tree. The decisions or final outcomes are represented by the leaves. And the data is separated at the decision nodes with Novel Healthcare.
Step 1: Define the issue in an organized manner..
Step 2: Create a model of the decision-making process.
Step 3: Use the right probability values and financial information.
Step 4: The decision tree must be "solved."
Step 5: Construct a sensitivity analysis.
Step 6: Make a list of your fundamental assumptions.

Statistical Analysis
The data for diabetic detection were collected from the cloud that contains over 60 participants in testing this system. The statistical software used for implementation in IBM SPSS version 21. The independent variables of the data are accuracy, Standard deviation and standard mean error and dependent variables in the data are Eye aspect ratio of x and y axis as parameters that is considered in the task. The independent sample T test analysis is carried out in this research work.

3. Results

The comparative analysis of the existing and the proposed algorithms is carried out by taking the accuracy rate of detection for both the algorithms. The diabetic prediction accuracy for both the algorithms is taken and can be used for analysis and comparison. The proposed algorithm is said to be more accurate than the existing algorithm ,and the accuracy rate of the proposed algorithm is found to be 86%.

Table 1 gives comparative study between the logistic regression algorithm and the Decision Tree algorithm.
Table 2 Group statistics T-Test for existing algorithm Mean and Decision Tree with the sample size 10. There is a statistically slight difference in the SD accuracy of the two algorithms, logistic regression(.06348)The Decision Tree algorithm had the highest accuracy (.07301).
Table 3 p-value where N=20 , significance value .000 from the statistical analysis tool.
Table 4 Comparison of the Accuracy of Decision Tree and logistic regression
Figure 1 represents the results are used as input into the statistical analysis tool and the graph is plotted using the values.

4. Discussion

The result shows that the decision tree with 86% accuracy appears to be better than the logistic regression algorithm with 79% accuracy (IEEE Staff 2020). The values of the Novel Effective framework are analyzed statistically and the difference is found out by plotting the graph against the algorithms. Diabetic prediction by remote monitoring is significantly efficient by monitoring the foot temperature which indicates foot ulcer caused by diabetes (Brooks et al. 2021).
Similar findings related to Decision Tree algorithm (Grąbczewski 2013)are significantly efficient in predicting the diabetic of the user compared to the existing algorithm that is the logistic regression algorithm ((Kaur, Sharma, and Sharma 2019; Barros, de Carvalho, and Freitas 2015)). The dataset containing a large number of images is given as input into both the algorithms, and the accuracy rate of prediction is obtained for the existing and the proposed algorithms (Deepak., John Justin Thangaraj, and Rajesh Khanna 2020). These values obtained are used for analysis and comparison for efficiency (Kaur, Sharma, and Sharma 2019; Barros, de Carvalho, and Freitas 2015).
The limitation is that developing securities issues requires a large team and support with cost effective measures so people can use this as a safety kit. In the future, the securities issues can be added (or) used in this project (Choi et al. 2019; Sosnovshchenko and Baiev 2018).

5. Conclusion

The research study found that the proposed novel Decision Tree algorithm is significantly efficient and accurate compared to the logistic regression algorithm. The accuracy of the diabetic prediction of the proposed algorithm is found to be 86%, and hence using the proposed algorithm gives better

results compared to the existing algorithm that has accuracy of 79%.

Declarations

Conflicts of Interest

No conflict of interest in this manuscript

Author Contribution

Author MNN is involved in data collection, data analysis and manuscript writing. Author VK was involved in conceptualization, data validation and critical review of the manuscript.

Acknowledgements

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding

Thankful for the following organizations for providing financial support that enabled us to complete the study.

1. Vee Eee Technologies Solution Pvt.Ltd., Chennai.
2. Saveetha University
3. Saveetha Institute of Medical and Technical Sciences
4. Saveetha School of Engineering.

6. Refrences

Baraneedharan, P., Sethumathavan Vadivel, C. A. Anil, S. Beer Mohamed, and Saravanan Rajendran. 2022. "Advances in Preparation, Mechanism and Applications of Various Carbon Materials in Environmental Applications: A Review." Chemosphere. https://doi.org/10.1016/j.chemosphere.2022.134596.

Barros, Rodrigo C., André C. P. L. F. de Carvalho, and Alex A. Freitas. 2015. Automatic Design of Decision-Tree Induction Algorithms. Springer.

Brooks, Elizabeth, Megan Burns, Ran Ma, Henk Jan Scholten, and Shawn Becker. 2021. "Remote Diabetic Foot Temperature Monitoring for Early Detection of Diabetic Foot Ulcers: A Cost-Effectiveness Analysis." ClinicoEconomics and Outcomes Research: CEOR 13 (October): 873–81.

Choi, Byoung Geol, Seung Woon Rha, Suhng Wook Kim, Jun Hyuk Kang, Ji Young Park, and Yung Kyun Noh. 2019. "Machine Learning for the Prediction of New-Onset Diabetes Mellitus during 5-Year Follow-up in Non-Diabetic Patients with Cardiovascular Risks." Yonsei Medical Journal 60 (2): 191–99.

Costache, Romulus, Alireza Arabameri, Thomas Blaschke, Quoc Bao Pham, Binh Thai Pham, Manish Pandey, Aman Arora, Nguyen Thi Thuy Linh, and Iulia Costache. 2021. "Flash-Flood Potential Mapping Using Deep Learning, Alternating Decision Trees and Data Provided by Remote Sensing Sensors." Sensors 21 (1). https://doi.org/10.3390/s21010280.

Deena, Santhana Raj, A. S. Vickram, S. Manikandan, R. Subbaiya, N. Karmegam, Balasubramani Ravindran, Soon Woong Chang, and Mukesh Kumar Awasthi. 2022. "Enhanced Biogas Production from Food Waste and Activated Sludge Using Advanced Techniques – A Review." Bioresource Technology. https://doi.org/10.1016/j.biortech.2022.127234.

Deepak., S. John Justin Thangaraj, and M. Rajesh Khanna. 2020. "An Improved Early Detection Method of Autism Spectrum Anarchy Using Euclidean Method." In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC). IEEE. https://doi.org/10.1109/i-smac49090.2020.9243361.

Ebben, K. C. W. J., M. P. Hendriks, L. Markus, M. Kos, I. H. J. T. de Hingh, J. R. Oddens, J. Rothbarth, et al. 2022. "Using Guideline-Based Clinical Decision Support in Oncological Multidisciplinary Team Meetings: A Prospective, Multicenter Concordance Study." International Journal for Quality in Health Care: Journal of the International Society for Quality in Health Care / ISQua, February. https://doi.org/10.1093/intqhc/mzac007.

Grąbczewski, Krzysztof. 2013. Meta-Learning in Decision Tree Induction. Springer.

IEEE Staff. 2020. 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS).

Karpagam, M., R. Beaulah Jeyavathana, Sathiya Kumar Chinnappan, K. V. Kanimozhi, and M. Sambath. 2022. "A Novel Face Recognition Model for Fighting against Human Trafficking in Surveillance Videos and Rescuing Victims." Soft Computing. https://doi.org/10.1007/s00500-022-06931-1.

Kaur, Gagandeep, Anshu Sharma, and Anurag Sharma. 2019. "Heart Disease Prediction Using KNN Classification Approach." International Journal of Computer Sciences and Engineering. https://doi.org/10.26438/ijcse/v7i5.416420.

Kumar, Amit, and Stefan Mozar. 2020. ICCCE

2020: Proceedings of the 3rd International Conference on Communications and Cyber Physical Engineering. Springer Nature.

Kumar, P. Ganesh, P. Ganesh Kumar, Rajendran Prabakaran, D. Sakthivadivel, P. Somasundaram, V. S. Vigneswaran, and Sung Chul Kim. 2022. "Ultrasonication Time Optimization for Multi-Walled Carbon Nanotube Based Therminol-55 Nanofluid: An Experimental Investigation." Journal of Thermal Analysis and Calorimetry. https://doi.org/10.1007/s10973-022-11298-4.

Nagarajan, Karthik, Arul Rajagopalan, S. Angalaeswari, L. Natrayan, and Wubishet Degife Mammo. 2022. "Combined Economic Emission Dispatch of Microgrid with the Incorporation of Renewable Energy Sources Using Improved Mayfly Optimization Algorithm." Computational Intelligence and Neuroscience 2022 (April): 6461690.

Nagaraju, V., B. R. Tapas Bapu, P. Bhuvaneswari, R. Anita, P. G. Kuppusamy, and S. Usha. 2022. "Role of Silicon Carbide Nanoparticle on Electromagnetic Interference Shielding Behavior of Carbon Fibre Epoxy Nanocomposites in 3-18GHz Frequency Bands." Silicon. https://doi.org/10.1007/s12633-022-01825-1.

Niu, Brenda, Narmin Mukhtarova, Oguzhan Alagoz, and Kara Hoppe. 2021. "Cost-Effectiveness of Telehealth with Remote Patient Monitoring for Postpartum Hypertension." The Journal of Maternal-Fetal & Neonatal Medicine: The Official Journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians, September, 1–7.

Pandiyan, P., R. Sitharthan, S. Saravanan, Natarajan Prabaharan, M. Ramji Tiwari, T. Chinnadurai, T. Yuvaraj, and K. R. Devabalaji. 2022. "A Comprehensive Review of the Prospects for Rural Electrification Using Stand-Alone and Hybrid Energy Technologies." Sustainable Energy Technologies and Assessments. https://doi.org/10.1016/j.seta.2022.102155.

Patel, Neil, Kathryn Kinmond, Pauline Jones, Pamela Birks, and Monica A. Spiteri. 2021. "Validation of COPDPredict™: Unique Combination of Remote Monitoring and Exacerbation Prediction to Support Preventative Management of COPD Exacerbations." International Journal of Chronic Obstructive Pulmonary Disease 16 (June): 1887–99.

Sosnovshchenko, Oleksandr, and Oleksandr Baiev. 2018. Machine Learning with Swift: Artificial Intelligence for iOS. Packt Publishing Ltd.

Sullivan, William. 2018. Decision Tree and Random Forest: Machine Learning and Algorithms: The Future Is Here! Createspace Independent Publishing Platform.

Tao, Xiaohui, Thanveer Basha Shaik, Niall Higgins, Raj Gururajan, and Xujuan Zhou. 2021. "Remote Patient Monitoring Using Radio Frequency Identification (RFID) Technology and Machine Learning for Early Detection of Suicidal Behaviour in Mental Health Facilities." Sensors 21 (3). https://doi.org/10.3390/s21030776.

Thenkabail, Prasad S. 2015. Land Resources Monitoring, Modeling, and Mapping with Remote Sensing. CRC Press.

Venu, Harish, Ibham Veza, Lokesh Selvam, Prabhu Appavu, V. Dhana Raju, Lingesan Subramani, and Jayashri N. Nair. 2022. "Analysis of Particle Size Diameter (PSD), Mass Fraction Burnt (MFB) and Particulate Number (PN) Emissions in a Diesel Engine Powered by Diesel/biodiesel/n-Amyl Alcohol Blends." Energy. https://doi.org/10.1016/j.energy.2022.123806.

Walton, Tracy. 2010. Medical Conditions and Massage Therapy: A Decision Tree Approach. Lippincott Williams & Wilkins.

Whangchai, Niwooti, Daovieng Yaibouathong, Pattranan Junluthin, Deepanraj Balakrishnan, Yuwalee Unpaprom, Rameshprabu Ramaraj, and Tipsukhon Pimpimol. 2022. "Effect of Biogas Sludge Meal Supplement in Feed on Growth Performance Molting Period and Production Cost of Giant Freshwater Prawn Culture." Chemosphere 301 (August): 134638.

Yaashikaa, P. R., M. Keerthana Devi, and P. Senthil Kumar. 2022. "Advances in the Application of Immobilized Enzyme for the Remediation of Hazardous Pollutant: A Review." Chemosphere 299 (July): 134390.

Tables and Figures

Table 1. Comparative study between the logistic regression algorithm and the Decision Tree algorithm with accuracy rate 89%.

S.No	logistic regression	Decision tree
1.	79.80	86.10
2.	79.88	86.15
3.	9.85	86.20
4.	79.93	86.25
5.	79.96	86.28
6.	79.97	86.32
7.	79.98	86.38
8.	79.99	86.42
9.	79.99	86.48
10.	80.00	86.53

Table 2. Group statistics T-Test for existing algorithm Mean (79.8840) and Decision Tree (86.1960) with the sample size 10. There is a statistically slight difference in the SD accuracy of the two algorithms. The Decision Tree algorithm had the highest accuracy and the (.07301) logistic regression(.06348).

Pair 1	**N**	**Mean**	**Std. deviation**	**Std.Error Mean**
Logistic regression	5	79.8840	.06348	.02839
Decision Tree	5	86.1960	.07301	.03265

Table 3. The correlation of the existing and the proposed algorithm with p-value where N=20 and significance value .000 from the statistical analysis tool.

Pair 1	N	Correlation	Significance value
logistic regression and Decision Tree	20	.997	.000

Table 4. Comparison of the F1 Score and Accuracy of Decision Tree and logistic regression. The independent sample effect sizes comparison of the existing and the proposed algorithms

	Levene's test for equality	**T- test for equality of means**

		of variances.								
		F	**Sig.**	t	**df**	**Sig.(2-tailed)**	**Mean difference**	**Std. error difference**	**95% confidence interval of the difference**	
									Lower	**Upper**
ACCURACY	Equal variance assumed	.136	.722	-145.886	8	.000	-6.31200	.04327	-6.41177	-6.21223
	Equal variances			-145.886	7.849	.000	-6.31200	.04327	-6.41211	-6.21189

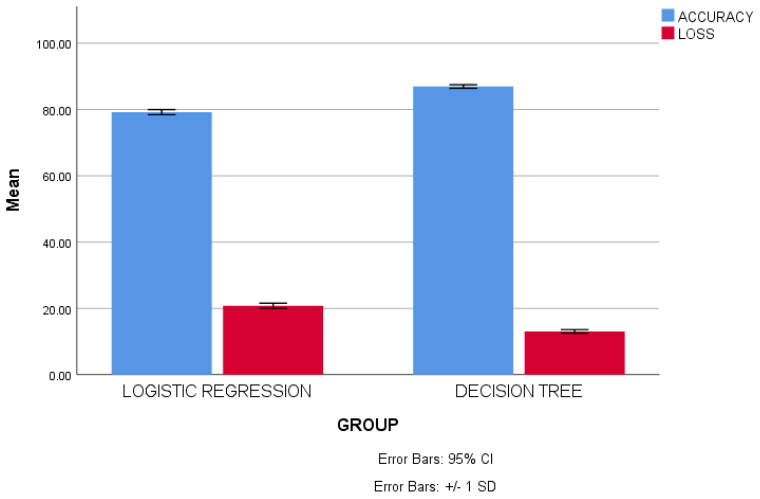

Fig. 1. Bar chart representation of the comparison of mean accuracy of the proposed and the existing algorithm. The accuracy of the prediction of the proposed algorithm is found to be79% and the proposed algorithm gives better results compared to the existing algorithm that has accuracy of 86% the mean accuracy detection is ±1 SD