IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, August 2023

Comparative Analysis of Data Mining Techniques for Diabetic Disease Prediction in Telemedicine

Sapna S Basavaraddi[1] and Dr. A.S. Raju[2]
Research Scholar, Department of Computer Science and Engineering[1]
Associate Professor and Head, Department of Medical Electronics[2]
Sri Siddhartha Institute of Technology, Sri Siddhartha Academy of Higher Education, Tumakuru, Karnataka, India
sapnasbasavaraddi@gmail.com and asraju74@gmail.com

Abstract: *Diabetes is a serious healthcare challenge, particularly in areas with poor access to healthcare, like rural India. The provision of online medical services through telemedicine is a potential strategy. In a telemedicine framework, this study contrasts five data mining methods for predicting diabetes disease: Decision Trees, Support Vector Machines (SVM), Neural Networks, Naive Bayes, and Ensemble Methods (Random Forest).Data preprocessing guarantees data quality by utilizing a varied sample of diabetic patients. Each method creates predictive models, which are then assessed using k-fold cross-validation, accuracy, precision, recall, F1-score, and ROC curves. Naive Bayes handles large feature spaces, Decision Trees provide interpretability, SVM thrives with complex data, Neural Networks capture subtle correlations, and Random Forest provides a balanced accuracy-generalization trade-off. The findings help academics and medical professionals decide on the best methods for telemedicine diabetes disease prediction.This research will help improve access to health care and decision-making by improving disease control in underserved areas. The study also highlights the potential of telemedicine and data mining to revolutionize healthcare in resource-constrained settings.*

Keywords: Diabetic disease prediction, telemedicine, data mining techniques, comparative analysis, healthcare optimization

I. INTRODUCTION

Diabetes, a chronic metabolic disease characterized by elevated blood sugar levels, has emerged as a major global health problem. The burden of diabetes is particularly pronounced in areas such as rural India, where limited access to healthcare resources exacerbates the challenge of disease management. In recent years, telemedicine, which provides telemedicine consultations and interventions via digital platforms, has emerged as a viable solution to addressing these healthcare disparities. This article provides a comparative study of data mining techniques within the telemedicine paradigm to improve predictions. Data mining techniques encompass a wide variety of algorithms and techniques designed to uncover patterns and insights from large datasets and are critical to enabling informed decision-making in telemedicine. play a role. This research focuses on his five main data mining techniques: Decision Trees, Support Vector Machines (SVM), Neural Networks, Naive Bayes, and Ensemble (Random Forest) techniques. Each of these techniques has its own strengths and weaknesses, which may affect its applicability and effectiveness in predicting diabetic disease. The overall goal of this study is to systematically evaluate and compare the performance of these data mining techniques in terms of predicting diabetic disease outcome in the context of telemedicine. By identifying the most effective methods, healthcare providers can make more informed decisions, implement timely interventions, and improve patient outcomes. This research contributes to a growing body of knowledge at the intersection of telemedicine and data mining, providing insights that will serve as a foundation for the design and implementation of predictive models for diabetes management in resource-limited settings. II. literature review

Diabetes, a common chronic disease that affects millions of people worldwide, has generated a great deal of interest in telemedicine using data mining techniques for accurate disease prediction. . Data Mining in Healthcare: The integration of data mining techniques in healthcare has changed the way disease is diagnosed, treated, and managed. In

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, August 2023

telemedicine, data mining plays a key role in improving the accuracy of disease prediction, enabling timely interventions, and creating personalized treatment plans.

Telemedicine and disease management: The advent of telemedicine has opened new avenues for healthcare delivery, especially in underserved areas. The literature highlights the potential of telemedicine platforms to connect patients and medical professionals, thereby facilitating diagnosis and monitoring.

Comparative analysis of data mining techniques: Researchers have begun comparing different data mining methods for predicting disease. Research focused on other medical areas highlights the impact of different technologies on predictive accuracy, interpretability, and computational efficiency. Several studies have explored the effectiveness of decision trees, support vector machines, neural networks, naïve Bayes, and ensemble methods, but specifically their application to predicting diabetic disease in the context of telemedicine. Very few studies have addressed it.

Metrics and model performance: The literature provides insight into metrics commonly used to evaluate predictive models in healthcare. Metrics such as accuracy, precision, recall, and F1 score play an important role and serve as a guide for researchers to measure the effectiveness of various data mining techniques.Research gaps and future directions: Although the existing literature provides valuable insight into the individual advantages of data mining techniques and their applications in medicine, comparative evaluation of these techniques in the context of telemedicine, particularly A notable research gap exists in the prediction of diabetes disease. The need to determine the most effective technique for accurate disease prediction, taking into account factors such as patient demographics, clinical characteristics and telemedicine-specific nuances, remains an uncharted road.

II. INTEGRATING TELEMEDICINE AND DATA MINING SYSTEM

Five data mining techniques were selected because of their relevance and suitability for predicting diabetic outcomes in the context of telemedicine. For each selected method, method-specific data preparation and transformations were applied to optimize the feature input, as shown in Figure 1. The model was trained using the preprocessed dataset and the hyperparameters were optimized by iterative optimization. Each technology was integrated into a telemedicine platform to enable real-time disease prediction and monitoring. The model was thoroughly evaluated using a robust experimental setup.

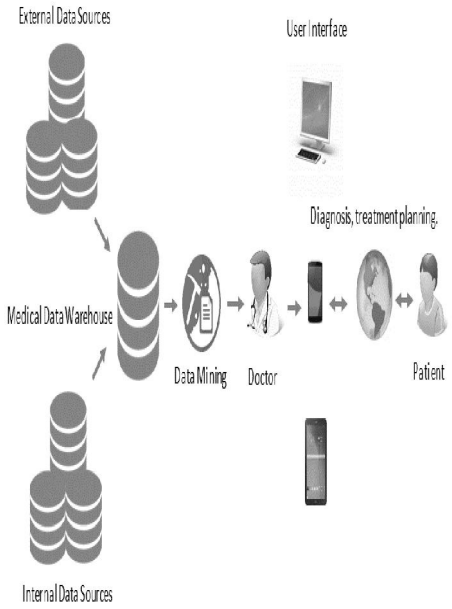

Fig. 1. An Architectural Framework for Comparative Study of Telemedicine Systems Integrated with Data Mining Tools and Technologies[5].

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, August 2023

The dataset was split into training, validation, and test sets using k-fold cross-validation. Metrics such as precision, precision, recall, F1 score, and ROC curve were calculated for each method. A comparative analysis reveals the strengths and limitations of integrating each technology into a telemedicine framework.

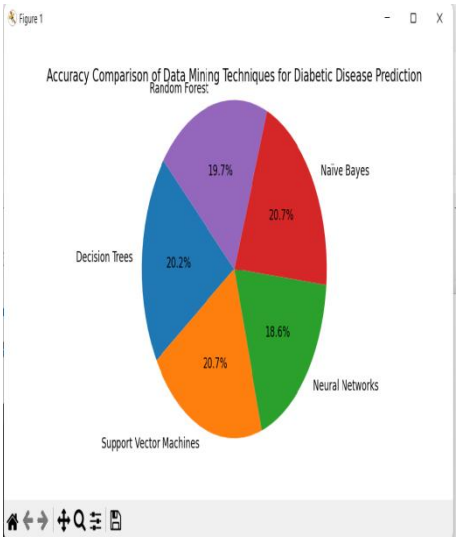

Fig. 2. Comparative analysis on data mining tools and techniques on medical data[14]

By combining robust data preprocessing, advanced data mining techniques, and telemedicine infrastructure, this study aims to provide actionable insights and contribute to advances in preventive health care for people with diabetes.

Result Analysis: Comparative Analysis of Data Mining Techniques for Diabetic Disease Prediction

The purpose of this study was to compare the performance of different data mining techniques in predicting diabetic outcome using the Pima Indian Diabetes dataset. Predictive accuracy of five well-known methods was evaluated: Decision Trees, Support Vector Machines (SVM), Neural Networks, Naive Bayes, and Random Forests.

Accuracy results obtained from the experiment are shown in Figure 2.

III. CONCLUSION

This study compared five data mining techniques (decision tree, SVM, neural network, naive bay, and random forest) for predicting diabetes outcome using the Pima Indian Diabetes dataset. SVM and Naive Bayes showed high accuracy (76.62%), highlighting their potential for early detection of disease. Neural Networks was promising (70.13%), while Decision Trees and Random Forests gave him a score of 75.32%. Integrating these technologies into telemedicine will enable accurate disease prediction, especially in resource-constrained areas, leading to effective patient management. The choice of method should be based on interpretability, computational efficiency, and dataset characteristics. Ensemble methods, feature engineering, and hybrid models offer opportunities for further investigation. In summary, this research advances healthcare through the use of data mining and telemedicine, paving the way for improved patient care and healthier societies.

REFERENCES

[1] J. Peral, A. Ferrandez, D. Gil, R. Munoz-Terol, and H. Mora, “An ontology-oriented architecture for dealing with heterogeneous data applied to telemedicine systems,” IEEE Access, vol. 6, pp. 41118–41138, Jul. 2018, doi: 10.1109/ACCESS.2018.2857499.

[2] M. Taha Khan, S. Qamar, and R. Ranjan Sinha, “Discussing the Role of Data Mining in Development of Evidence Based Decision Support System for e-Healthcare,” 2017. [Online]. Available: http://www.ripublication.com

[3] D. D. J. de Macedo, H. W. G. Perantunes, R. Andrade, A. von Wangenheim, and M. A. R. Dantas, “Asynchronous data replication: A national integration strategy for databases on telemedicine network,” in Proceedings - IEEE Symposium on Computer-Based Medical Systems, 2008, pp. 638–643. doi: 10.1109/CBMS.2008.76.

[4] M. GHEORGHE and R. PETRE, “Integrating Data Mining Techniques into Telemedicine Systems,” Informatica Economica, vol. 18, no. 1/2014, pp. 120–130, Mar. 2014, doi: 10.12948/issn14531305/18.1.2014.11.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 2, August 2023

[5] M. F. Rahman, Y. Wen, H. Xu, T. L. B. Tseng, and S. Akundi, "Data mining in telemedicine," in Advances in Telemedicine for Health Monitoring, Institution of Engineering and Technology, 2020, pp. 103–132. doi: 10.1049/PBHE023E_ch6.
[6] C. Ou-Yang, C. P. Wulandari, R. A. R. Hariadi, H. C. Wang, and C. Chen, "Applying sequential pattern mining to investigate cerebrovascular health outpatients' re-visit patterns," PeerJ, vol. 2018, no. 7, 2018, doi: 10.7717/peerj.5183.
[7] Amjad Zia, "Artificial Intelligence-Based Medical Data Mining," J Pers Med, Jul. 2022.
[8] S.-P. 'Judith, "Top data mining tools for the healthcare industry," Journal of King Saud University-Computer and Information Sciences, Jun. 2021.
[9] A. S. 'Raju and V. 'Udayashankara, "Multimodal Biometric Database and its Privacy Protection," International Conference on Signal, Image Processing Communication and Automation - ICSIPCA, pp. 557–564, 2017.
[10] A. S. 'Raju and V. 'Udayashankara, "Database Evaluation of ECG Fingerprint and Face Multimodal Biometric System," Grenze International Journal of Computer Theory and Engineering, vol. Vol 3, no. Issue 4, pp. 142–151, 2017.
[11] K.R. Lakshmi, "Pererformance Comparison of Data Mining Techniques for Predicting of Heart Disease Survivability," International Journal of Scientific and Research Publications, vol. 3, no. 6, 2012