

Intelligent Health Care System Online on Diabetes Prediction Using SVM and Naive Bayes

Ramkumar, Manoj Tiwari

Assistant Professor, Sahyadri Valley College of Engineering & Technology, Rajuri

Article Information

Received : 10 Feb 2025
Revised : 16 Feb 2025
Accepted : 20 Feb 2025
Published : 25 Feb 2025

Corresponding Author:

Ramkumar

Abstract— This paper presents Diabetes is common problem over the world. Many of us are facing this disease. In this disease, our blood sugar level get disturbed. So we are trying to build a tool that can help people to check whether they are suffering from it or not? It might have happened so many times that you or someone yours need doctors help immediately , but they are not available due to some reason. The health prediction system is an end user support and online consultation project. Here we propose a system that allows users to get instant guidance on their health issues through an intelligent health care system online. The system is fed with various symptoms and the disease associated with those systems. Diabetes is considered as one of the deadliest and chronic diseases which causes an increase in blood sugar. Many complications occur if diabetes remains untreated and unidentified. The tedious identifying process results in visiting of a patient to a diagnostic center and consulting doctor. But the rise in machine learning approaches solves this critical problem. The motive of this study is to design a model which can prognosticate the likelihood of diabetes in patients with maximum accuracy. Therefore three machine learning classification algorithms namely Decision Tree, SVM and Naive Bayes are used in this experiment to detect diabetes at an early stage. Experiments are performed on Pima Indians Diabetes Database (PIDD) which is sourced from UCI machine learning repository. The performances of all the three algorithms are evaluated on various measures like Precision, Accuracy, F-Measure, and Recall.

Keywords: *Diabetes, Mining, Data, SVM, Decision tree*

Copyright © 2025: Ramkumar, Manoj Tiwari, This is an open access distribution, and reproduction in any medium, provided Access article distributed under the Creative Commons Attribution License the original work is properly cited License, which permits unrestricted use.

Citation: Ramkumar, Manoj Tiwari, “Intelligent Health Care System Online on Diabetes Prediction Using SVM and Naive Bayes”, Journal of Science, Computing and Engineering Research, 8(2), February 2025.

I. INTRODUCTION

Diabetes mellitus is one of the world's major diseases. Millions of people are affected by the disease. The risk of diabetes is increasing day by day and is found mostly in women than men. The diagnosis of diabetes is a tedious process. So with improvement in science and technology it is made easy to predict the disease.

The purpose is to diagnose whether the person is affected by diabetes or not using K Nearest Neighbor classification technique. The diabetes dataset is taken as the training data and the details of the patient are taken as testing data. Data mining is a subfield in the subject of software engineering. It is the methodical procedure of finding examples in huge data sets including techniques at the crossing point of manufactured intelligence, machine learning, insights, and database systems.

The goal of the data mining methodology is to think data from a data set and change it into a reasonable structure for further use. Our examination concentrates on this part of Medical conclusion learning design through the gathered

data of diabetes and to create smart therapeutic choice emotionally supportive network to help the physicians. Data mining is a significant tool in medical databases, which enhances the sensitivity and/or specificity of disease detection and diagnosis by opening a window of relatively better resources [4].

Applying machine learning and data mining methods in diabetes research is a pivotal way to utilizing plentiful available diabetes-related data for extracting knowledge. The severe social impact of the specific disease makes DM one of the main priorities in medical science research, which inevitably produces large amounts of data.

Therefore, there is no doubt that machine learning and data mining approaches in DM are of great concern on diagnosis, management, and other related clinical administration aspects [5].

In order to achieve the best classification accuracy, abundant algorithms and diverse approaches have been applied, such as traditional machine learning algorithms, ensemble learning approaches, and association rule learning.

Most noted among the aforementioned ones are the following: Calisir and Dogantekin proposed LDA-MWSVM, a system for diabetes diagnosis [6].

The system performs feature extraction and reduction using the Linear Discriminant Analysis (LDA) method, followed by classification using the Morlet Wavelet Support Vector Machine (MWSVM) classifier. Gangji and Abadeh [7] presented an Ant Colonybased classification system to extract a set of fuzzy rules, named FCSANTMINER, for diabetes diagnosis. In [8], authors regard glucose prediction as a multivariate regression problem utilizing Support Vector Regression (SVR).

Agarwal [9] utilized semi-automatically marked training sets to create phenotype models via machine learning methods. Ensemble approaches, which utilize multiple learning algorithms, have been confirmed to be an effective way of enhancing classification accuracy.

II. LITERATURE SURVEY

The health industry has been growing a lot from past few years. This technique has gained a lot of importance in medical areas. It has been calculated that a care hospital may generate five terabytes of data in the year.

In our day to day life we have lot of other problems to deal with and we neglect our health problems. So in order to overcome such problem we have designed user friendly website which helps users to get diagnosed from their residence at any time. We also provide an option for booking an appointment with the doctor to discuss health related problems and get diagnosed properly.

Sajida et al. in [2] discusses the role of Adaboost and Bagging ensemble machine learning methods [8] using J48 decision tree as the basis for classifying the Diabetes Mellitus and patients as diabetic or non diabetic, based on diabetes risk factors. Results achieved after the experiment proves that, Adaboost machine learning ensemble technique outperforms well comparatively bagging as well as a J48 decision tree.

Orabi et al. in [9] designed a system for diabetes prediction, whose main aim is the prediction of diabetes a candidate is suffering at a particular age. The proposed system is designed based on the concept of machine learning, by applying decision tree. Obtained results were satisfactory as the designed system works well in predicting the diabetes incidents at a particular age, with higher accuracy using Decision tree, [7].

Pradhan et al in [4] used Genetic programming (GP) for the training and testing of the database for prediction of diabetes by employing Diabetes data set which is sourced from UCI

repository. Results achieved using Genetic Programming [5],

It gives optimal accuracy as compared to other implemented techniques. There can be significant improve in accuracy by taking less time for classifier generation. It proves to be useful for diabetes prediction at low cost.

Rashid et al. in [8] designed a prediction model with two sub-modules to predict diabetes schronic disease. ANN (Artificial Neural Network) is used in the first module and FBS (Fasting Blood Sugar) is used in the second module. Decision Tree (DT)[1] is used to detect the symptoms of diabetes on patient's health.

Nongyao et al. in [7] applied an algorithm which classifies the risk of diabetes mellitus. To fulfill the objective author has employed four following renowned machine learning classification methods namely Decision Tree, Artificial Neural Networks, Logistic Regression and Naive Bayes. For improving the robustness of designed model Bagging and Boosting techniques are used. Experimentation results shows the Random Forest algorithm gives optimum results among all the algorithms employed.

III. VARIOUS ISSUES OF DIABETES PREDICTION IN DATA MINING

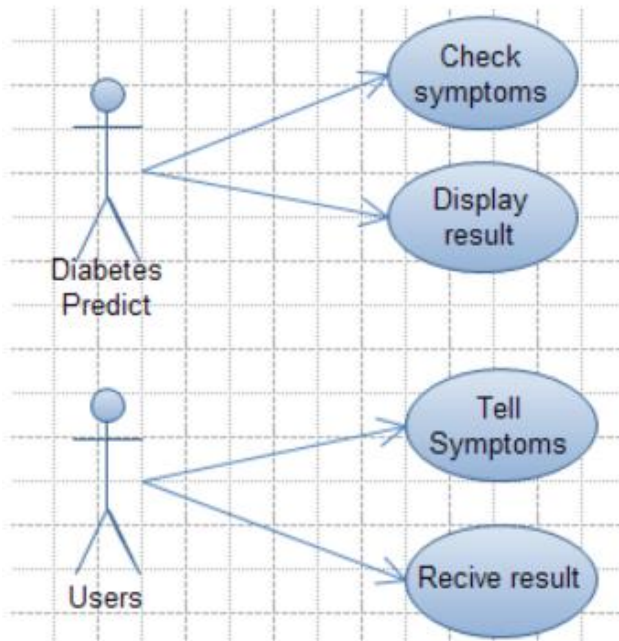
Diabetes is a chronic disease characterized by a long treatment cycle, numerous complications (e.g., kidney and eye diseases), and recurrent illness. With advances in the informatization of medicine, medical industries with large amounts of complicated patient data are keen to extract information from this data to assist the development of these industries.

Simultaneously, they also seek to be capable of alleviating the challenges faced by medical personnel, through the forthcoming development of smart medicine. The use of machine learning and other artificial intelligence methods for the analysis of medical data in order to assist diagnosis and treatment is one of the manifestations of smart medicine with the most practical significance.

With the improvement of the living standards of our people and the westernization of our diet, the incidence, mortality, and morbidity of diabetes have significantly increased and have a serious impact on our health. In 2006, Shang [1] made use of the survey data of Xinjiang chronic disease integrated prevention and control demonstration site in the New Urban District of Urumqi in 2004 and surveyed 2031 people over the age of 18 in three communities in the district.

The results showed the relationship between diabetes and age and gender: the prevalence of male and female rose with age, because the decrease of glucose tolerance with age and the improvement of living standard are the reasons for the increased incidence. Overweight and obesity are one of the

risk factors of diabetes mellitus. The survey found that the prevalence of diabetes in people with BMI>24 was 10.58%, the prevalence of diabetes in people with BMI≤24 was 4.31%, two groups prevalence by chi-square test was P



IV. CONCLUSION

One of the important real-world medical problems is the detection of diabetes at its early stage. In this study, systematic efforts are made in designing a system which results in the prediction of disease like diabetes. During this work, three machine learning classification algorithms are studied and evaluated on various measures. Experiments are performed on Pima Indians Diabetes Database. In future, the designed system with the used machine learning classification algorithms can be used to predict or diagnose other diseases

REFERENCES

- [1]. P. Nirmala, T. Manimegalai, J. R. Arunkumar, S. Vimala, G. Vinoth Rajkumar, Raja Raju, "A Mechanism for Detecting the Intruder in the Network through a Stacking Dilated CNN Model", *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 1955009, 13 pages, 2022. <https://doi.org/10.1155/2022/1955009>.
- [2]. D. Sathyanarayanan, T. S. Reddy, A. Sathish, P. Geetha, J. R. Arunkumar and S. P. K. Deepak, "American Sign Language Recognition System for Numerical and Alphabets," 2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE), Chennai, India, 2023, pp. 1-6, doi: 10.1109/RMKMATE59243.2023.10369455.
- [3]. J. R. Arunkumar, Tagele berihun Mengist, 2020" Developing Ethiopian Yirgacheffe Coffee Grading Model using a Deep Learning Classifier" *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-9 Issue-4, February 2020. DOI: 10.35940/ijitee.D1823.029420.
- [4]. Ashwini, S., Arunkumar, J.R., Prabu, R.T. et al. Diagnosis and multi-classification of lung diseases in CXR images using optimized deep convolutional neural network. *Soft Comput* (2023). <https://doi.org/10.1007/s00500-023-09480-3>
- [5]. J.R.Arunkumar, Dr.E.Muthukumar," A Novel Method to Improve AODV Protocol for WSN" in *Journal of Engineering Sciences* ISSN NO: 0377-9254 Volume 3, Issue 1, Jul 2012.
- [6]. R. K, A. Shameem, P. Biswas, B. T. Geetha, J. R. Arunkumar and P. K. Lakineni, "Supply Chain Management Using Blockchain: Opportunities, Challenges, and Future Directions," 2023 Second International Conference on Informatics (ICI), Noida, India, 2023, pp. 1-6, doi: 10.1109/ICI60088.2023.10421633.
- [7]. Arunkumar, J. R. "Study Analysis of Cloud Security Chanllenges and Issues in Cloud Computing Technologies." *Journal of Science, Computing and Engineering Research* 6.8 (2023): 06-10.
- [8]. J. R. Arunkumar, R. Raman, S. Sivakumar and R. Pavithra, "Wearable Devices for Patient Monitoring System using IoT," 2023 8th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2023, pp. 381-385, doi: 10.1109/ICCES57224.2023.10192741.
- [9]. S. Sugumaran, C. Geetha, S. S, P. C. Bharath Kumar, T. D. Subha and J. R. Arunkumar, "Energy Efficient Routing Algorithm with Mobile Sink Assistance in Wireless Sensor Networks," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ACCAI58221.2023.10201142.
- [10]. R. S. Vignesh, V. Chinnammal, Gururaj.D, A. K. Kumar, K. V. Karthikeyan and J. R. Arunkumar, "Secured Data Access and Control Abilities Management over Cloud Environment using Novel Cryptographic Principles," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-8, doi: 10.1109/ACCAI58221.2023.10199616.
- [11]. Syamala, M., Anusuya, R., Sonkar, S.K. et al. Big data analytics for dynamic network slicing in 5G and beyond with dynamic user preferences. *Opt Quant Electron* 56, 61 (2024). <https://doi.org/10.1007/s11082-023-05663-2>
- [12]. Krishna Veni, S. R., and R. Anusuya. "Design and Study Analysis Automated Recognition system of Fake Currency Notes." *Journal of Science, Computing and Engineering Research* 6.6 (2023): 16-20.
- [13]. V. RamKumar, S. Shanthi, K. S. Kumar, S. Kanageswari, S. Mahalakshmi and R. Anusuya, "Internet of Things Assisted Remote Health and Safety Monitoring Scheme Using Intelligent Sensors," 2023 International Conference on Advances in Computing, Communication and Applied

- Informatics (ACCAI), Chennai, India, 2023, pp. 1-8, doi: 10.1109/ACCAI58221.2023.10199766.
- [14]. R. S. Vignesh, R. Sankar, A. Balaji, K. S. Kumar, V. Sharmila Bhargavi and R. Anusuya, "IoT Assisted Drunk and Drive People Identification to Avoid Accidents and Ensure Road Safety Measures," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ACCAI58221.2023.10200809.
- [15]. I. Chandra, G. Sowmiya, G. Charulatha, S. D, S. Gomathi and R. Anusuya, "An efficient Intelligent Systems for Low-Power Consumption Zigbee-Based Wearable Device for Voice Data Transmission," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10083856.
- [16]. G. Karthikeyan, D. T. G, R. Anusuya, K. K. G, J. T and R. T. Prabu, "Real-Time Sidewalk Crack Identification and Classification based on Convolutional Neural Network using Thermal Images," 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 2022, pp. 1266-1274, doi: 10.1109/ICACRS55517.2022.10029202.
- [17]. R. Meena, T. Kavitha, A. K. S, D. M. Mathew, R. Anusuya and G. Karthik, "Extracting Behavioral Characteristics of College Students Using Data Mining on Big Data," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10084276.
- [18]. S. Bharathi, A. Balaji, D. Irene, J. C. Kalaivanan and R. Anusuya, "An Efficient Liver Disease Prediction based on Deep Convolutional Neural Network using Biopsy Images," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022, pp. 1141-1147, doi: 10.1109/ICOSEC54921.2022.9951870.
- [19]. I. Chandra, G. Sowmiya, G. Charulatha, S. D, S. Gomathi and R. Anusuya, "An efficient Intelligent Systems for Low-Power Consumption Zigbee-Based Wearable Device for Voice Data Transmission," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ICECONF57129.2023.10083856.
- [20]. Revathi, S., et al. "Developing an Infant Monitoring System using IoT (INMOS)." International Scientific Journal of Contemporary Research in Engineering Science and Management 6.1 (2021): 111-115.
- [21]. R. K, A. Shameem, P. Biswas, B. T. Geetha, J. R. Arunkumar and P. K. Lakshini, "Supply Chain Management Using Blockchain: Opportunities, Challenges, and Future Directions," 2023 Second International Conference on Informatics (ICI), Noida, India, 2023, pp. 1-6, doi: 10.1109/ICI60088.2023.10421633.
- [22]. J.R.Arunkumar. "Comprehensive Analysis of Security Issues in Cloud Computing Technologies", Journal of Science, Computing and Engineering Research, 6(5), 06-10, June 2023.
- [23]. S. Sugumaran, C. Geetha, S. S, P. C. Bharath Kumar, T. D. Subha and J. R. Arunkumar, "Energy Efficient Routing Algorithm with Mobile Sink Assistance in Wireless Sensor Networks," 2023 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), Chennai, India, 2023, pp. 1-7, doi: 10.1109/ACCAI58221.2023.10201142.
- [24]. I. Chandra, K. V. Karthikeyan, R. V, S. K, M. Tamilselvi and J. R. Arunkumar, "A Robust and Efficient Computational Offloading and Task Scheduling Model in Mobile Cloud Computing," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-8, doi: 10.1109/ICECONF57129.2023.10084293.
- [25]. R. S. Vignesh, A. Kumar S, T. M. Amirthalakshmi, P. Delphy, J. R. Arunkumar and S. Kamatchi, "An Efficient and Intelligent Systems for Internet of Things Based Health Observance System for Covid 19 Patients," 2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF), Chennai, India, 2023, pp. 1-8, doi: 10.1109/ICECONF57129.2023.10084066.
- [26]. DC Jullie Josephine, J Sudhakar, T Helan Vidhya, R Anusuya, G Ramkumar, "An Improved Multi class Breast cancer classification and Abnormality Detection based on Modified Deep Learning Neural Network Principles", Deep Learning in Biomedical Signal and Medical Imaging, CRC Press, Taylor and Francis, 2024.
- [27]. R. Anusuya, Pragya Vashishtha, "Real Automatic Number Plate Image Detection With Yolo Algorithms", Journal of Science, Computing and Engineering Research, 7(7), July 2024.