Classification of Diabetes Disease using Adaptive Bio-Inspired Gene-Level Deep Neural Networks

Jyothsna Devi.Kuchipudi[1], Dr.B.Srilatha[2], Dr. Rajkumar Gangappa Nadakinamani[3], Dr. Vani V G[4], Dr. Rakhi Kamra[5], Dillip Narayan Sahu 6*

[1]Assistant Professor, Department of Food Technology, Sree Sastha Institute of Engineering and Technology, Chembarambakkam, Chennai, India.

[2]Associate Professor, Department of Electronics and Communication Engineering, The Oxford College of Engineering, India.

[3]Specialist Cardiologist, Badr Al samaa Hospital, Muscat.

[4]Assistant Professor, Department of Computer Science and Engineering, Government Engineering College, Hassan Karnataka, India.

[5]Assistant Professor, Department of Electrical and Electronics Engineering, Maharaja Surajmal Institute of Technology , New Delhi, Delhi, India.

[6*]Assistant Professor, Department of MCA, Gangadhar Meher University, Odisha, India.

*Email: dillip1seminar@gmail.com

Abstract

The vast amount of information in a medical database makes data classification a difficult challenge in data mining. When used to medical data, associative classification improves classification accuracy and disease prediction. With this goal in mind, the proposed study presents three methods for accurately categorizing patient medical data via the creation of optimal association rules. The three methods are Logistic Fully Recurrent Deep Neural Learning Classification (LFRDNLC), Adaptive Bio-Inspired Gene Optimization Based Deep Neural Associative Classification (ABGO-DNAC), and Gene Optimized Association Rule Generation based Integral Derivative Gradient Boost Classification (GOARG-IDGBC). The improved results for diabetic illness diagnosis with higher classification accuracy and less time consumption are produced by using the aforementioned three recommended methodologies. The suggested GOARG-IDGBC method's primary objective is to boost classification precision while diagnosing diabetes. The fitness function of each characteristic is evaluated as part of an optimized evolutionary algorithm to generate the best possible set of association rules. The suggested GOARG-IDGBC method employs an integral derivative gradient boost classifier to perform classification based on previously specified association rules (IDGBC). Attributes are categorized using a decision tree in IDGBC.

Keywords: Data mining, Medical data, LFRDNLC, Deep Neural Learning, Gradient and Gene Optimization.

1. Introduction

Diabetes is a lifelong condition that poses serious danger to human health. Diabetes is characterized by a blood glucose level that persists over the usual range. Inadequate insulin secretion or the negative physiologic consequences of insulin produce this aberrant state. Chronic damage and dysfunction of many tissues, including the eyes, kidneys, heart, blood vessels, and nerves, are also caused by diabetes. Diabetes is becoming increasingly common as people's level of living improves. As a result, it's more important than ever to study how to properly diagnose diabetes. Fasting blood glucose, glucose tolerance, and random blood glucose levels are used to make a diabetes diagnosis in the medical field. Diagnosing diabetes early allows for simple management[1]. Traditional methods of diagnosis failed because of their unreliable prognostic findings. In order to extract useful information from large medical datasets, association discovery is a frequent data mining technique. Finding the meaningful connections between diabetes patient data and healthy patient data is made possible by association discovery[2]. Data mining techniques are widely employed in the diagnosis of diabetes, with positive outcomes. Using information gleaned from patients' routine medical checks, data mining can help patients make an initial diagnosis of diabetes mellitus. In data mining methods, the most crucial challenges involve picking the right characteristics to mine and a reliable classifier to apply to those features. The goal of pattern classification is to establish a model using training data that can correctly identify the category of test patterns. There have been attempts to merge the benefits of classification with those of association discovery. Diabetic diagnosis is greatly aided by the use of classification in conjunction with association rule generation[3]. When compared to other classification methods, neural networks have shown to be effective in predicting diabetes. Diseases like diabetes mellitus are commonplace. More than 415 million people worldwide have diabetes, according to the World Diabetes Federation. Diabetes is a metabolic illness in which the body is unable to produce or respond appropriately to insulin, a hormone necessary for maintaining normal glucose levels in the blood. Type 1 diabetes is characterized by a complete lack of insulin due to the death of beta cells in the pancreas by the body's own immune system. Insulin resistance and a little insulin shortage characterize the more common type 2 diabetes.

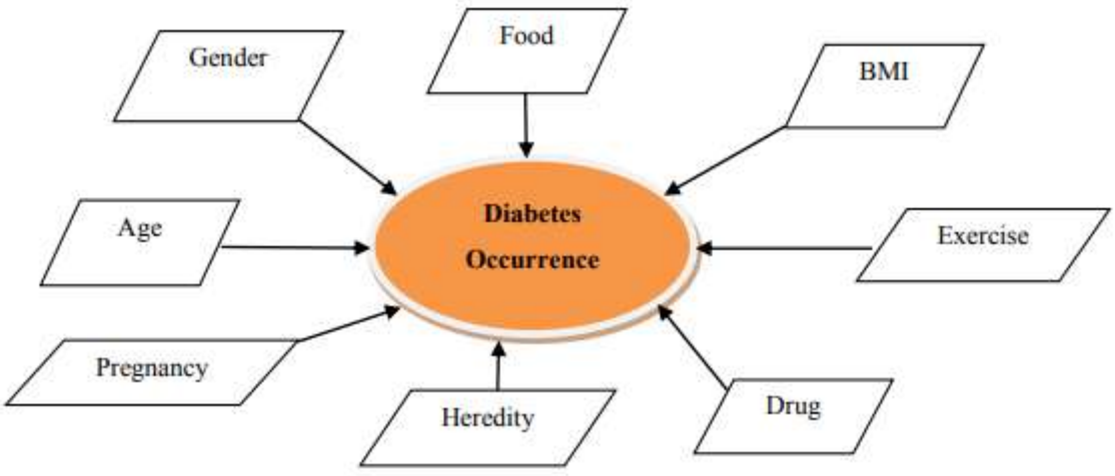

Figure .1: Factors Related to Diabetic Occurrence in Humans

Dietary shifts are largely to blame for the rapid global spread of diabetes, as reported by the World Development Indicators (WDI). The risk of other forms of organ and tissue dysfunction is likewise raised by diabetes[4]. Consequently, it is crucial in modern medicine to detect diabetes early. Yet, the root of this illness has yet to be determined. Diabetic disease is likely caused by a combination of genetic predisposition, environmental influences, and personal choices. The risk factors for developing diabetes in humans are displayed in Figure.1. The problem of early diabetes detection persists despite the disease's rising prevalence due to the limitations of current diagnostic tools. Inaccurate or delayed diagnosis is another key problem that hinders treatment for this condition. As a result, a great deal of work is needed to establish a more efficient technique of diabetic diagnosis to manage the disease at an earlier stage. Diagnosing diabetes can be challenging for doctors due to the number of criteria that must be considered. A doctor makes judgments by comparing current test results with past judgments made about the same patient under comparable circumstances[5]. Nonetheless, taking into account such a large number of parameters makes this type of diagnostic a breeze. In addition, there is an abundance of data regarding patients, illnesses, and doctors in medical datasets. It takes a battery of pricey tests to properly diagnose a condition. Using data mining techniques decreases the expense associated with disease forecasting and diagnosis. Hence, data mining techniques are utilized to discover crucial principles for efficient detection of diabetes and its consequences in order to diagnose the disease and classify patients. Mining large medical datasets for hidden patterns is one application of medical data mining. Diseases are diagnosed using data analysis results.

2. Literature Survey

Large amounts of data regarding patients, diseases, and doctors are used in medical databases. Several costly tests were needed to diagnose the disease. Using machine learning and data mining techniques reduces the financial burden associated with disease prognosis and diagnosis. Because of the rise in obesity and the decline in physical activity, diabetes is often referred to be a disease of modern civilization. If diabetes is not properly identified and treated in a timely manner, it can lead to significant complications and even death. Complications from diabetes can be avoided with early detection and diagnosis. Diabetes mellitus refers to a range of metabolic illnesses characterized by persistently elevated blood glucose levels and is the more common term. This happens when either not enough insulin is produced or when cells in the body do not respond normally to insulin. The process of classification is a data mining procedure that involves sorting data into categories. Classifying data is used in many scientific and engineering disciplines to organize data into more manageable groups. The diagnosis of health care data is a crucial process that requires a high level of expertise and experience. Classification of diabetes data is a crucial step in making a diagnosis. In order to diagnose a problem, it is necessary to analyze the data and extract useful information from it. There have been many attempts to use symptoms to diagnose diabetes, however this approach has led to faulty assumptions and has not yielded quality criteria for building an effective associative classifier. Hence, we investigate the use of association rule mining with optimization and classification strategies to improve the diagnostic accuracy for diabetes. Ant colony based categorization method is discussed in [6] for capturing a set of fuzzy rules for diabetes diagnosis. The diagnosis of diabetes was improved with the use of Ant Colony Optimization and Fuzzy Logic. Prediction of diabetic disease was improved by using an interpretable fuzzy rule base, which led to a greater level of classification accuracy. This ultimately helped implement the best fuzzy rule base for diabetes diagnosis that had been discovered. Yet, little effort has been made to reduce the computing cost of Ant Colony-based categorization systems. Ant colony optimization (ACO) and genetic programming (GP) were combined to form a hybrid metaheuristic in [7]. (GA). During the ACO phase, we used GA's population to power decision lists fed by a wide variety of data subsets. The classification accuracy of the decision lists was enhanced using GA, and the model size was reduced. Unfortunately, utilizing a hybrid metaheuristic did not overcome the issue of time complexity. In [8], the authors presented a novel method for medical diagnosis that combined genetic algorithms (GAs) and particle swarm optimization (PSO) (GAPSO-FS). SVM parameters were adjusted using swarm optimization techniques, and the best feature subset was chosen using a Genetic Algorithm (GA). There are three levels total in GAPSO-FS: an optimization layer, an intermediate layer, and a final layer. Data from two optimization levels was harmonized at the intermediate layer, and the processed data was distributed to the various tiers. PSO helped find the optimal combination of continuous parameters. When tasked with selecting the best SVM model parameters and feature subset for SVM classifier, GAPSO-FS excelled. As a result, the accuracy of medical diagnoses was improved by classification. Yet, multiclass concerns in practical contexts were not effectively handled. To improve classification accuracy, the authors of [9] discussed an ant colony optimization based on associations. Attribute-level data set refinement was used in this optimization strategy. The final dataset was collected, and the support value was used to calculate an estimate of the maximum threshold. These support values were collected, and using them, we determined whether or not the dataset was useful. In the end, the ACO approach was employed to determine the outcomes with greater precision in classification. In [10], the authors created a hybrid model for classifying medical data by combining a case-based reasoning strategy with a Particle Swarm Optimization (PSO) model. To preprocess the data set and establish a weight vector for each characteristic, a case-based reasoning strategy was used. Then, a disease recognition decision-making system was built using a particle swarm optimization model. By using PSO to limit the impact of starting points, the dataset was partitioned into a sizable number of clusters, the size of which was later reduced. It helped produce results that medical professionals could understand and use in their diagnoses. Medical data categorization may have been improved with the use of additional soft computing approaches, however this was not done. The supervised particle swarm optimization (S-PSO) classification approach to defect identification is explored by the authors in [11]. The effects of both the local and global search were stabilized by the development of a particle position updating method that combines a fixed iteration interval intervention updating strategy. As a result, there was a greater capability to search for and guide the particles, as well as a greater diversity of particles. As a result, an estimated fitness function was used to limit the output to class centers that were optimal in terms of intra-class distance. The classification accuracy was improved by using the S-PSO algorithm. S-PSO was unable to produce reliable forecasts because of its inefficient performance. In order to improve classification accuracy while keeping costs down, the authors of [12] address fuzzy and genetic approaches to identifying the existence of diabetes. The genetic technique was used to execute the idea of feature selection. This ultimately helped with diabetes diagnosis. Using membership functions, we performed a simultaneous mapping that relied on an appropriateness measure of variable values to each class. After that, we used fuzzy thinking techniques to improve our categorization accuracy. Fuzzy and genetic approaches were used to get effective categorization results. Nonetheless, the procedure for selecting important traits was unsuccessful. The hybrid strategy proposed by the author in [13] makes use of both the artificial bee colony (ABC) method and support vector machines to choose features and classify data, respectively. For effective classification using the SVM classifier, the datasets had their irrelevant and obsolete features removed. Liver disease and diabetes were two other areas where the hybrid method was used for diagnosis. There was a significant improvement in classification accuracy, and the method is still usable in pattern recognition. An automated system was built to combine the rules and statistical models presented in [14], which used rule-based classification to aid in the diagnosis of malaria. Clinical diagnosis of malaria was modeled after a statistical prototype. Overall classification rule sensitivity was increased, and data mining classification rule performance was enhanced. Rule-based classification, however, did not improve accuracy. In [15], the authors evaluated a CAD for early Alzheimer's disease diagnosis that combined continuous attribute discretization with association rule mining (AD). Emission evaluated tomography pictures were required for this. Using picture histogram segmentation, we were able to extract a mask from the average of the control photos. The 3D voxels centered in mask coordinates were then selected using a discretization method based on the mean intensity and equal-width binning. These ROIs, along with photos of standard subjects intended to represent the expected image pattern, were employed as input for an Association Rule (AR)-mining system. By establishing minimal support and confidence levels at each discretization level, rules with more predictive ability could be constructed. Finally, we used classification to evaluate how many ARs were confirmed by each participant. A hybrid classifier for diabetes was published by the authors in [16]; it is called the Logistic Adaptive Network-based Fuzzy Inference System (LANFIS). Logistic regression and the Adaptive Network-based Fuzzy Inference System were combined to form LANFIS. Classifiers were discouraged as a means of attaining continuous output since they do not make use of redundant attributes and so increase the number of tests performed during data collecting. Using the application of a subtractive clustering method to generate fuzzy rules and the elimination of missing values, the LANFIS intelligent system was able to outperform competing fuzzy classifiers. Unfortunately, the hybrid classifier did not significantly improve classification accuracy. For diabetes diagnosis, the authors of [17] created a Reinforcement Learning-based Evolutionary Fuzzy Rule-Based System (RLEFRBS). Using numerical data, a Rule Base (RB) was built, and rules were optimized. Depending on the level of confidence, extraneous rules were eliminated, and criteria from precursor portions were taken to provide for more easily understandable rules. After that, a

GA was used to pick the best set of rules to use. Reinforcement learning was used in conjunction with evolutionary tuning of membership functions and weight modifying to further improve RLEFRBS' performance (RL). In addition, revealed occurrences were dealt with using a rule stretching technique that proved effective. The difficulties of medical diagnosis prompted the creation of the Neighborhood Rough Set Based Classification (NRSC) algorithm. The NRSC algorithm assisted in the classification of healthcare data. With its help, the doctors were able to make more informed decisions and the categorization performance improved. Yet, classification's effectiveness in disease prediction was lacking. Based on Rough Set Theory (RST) and the Bat optimization Algorithm (BA), RST-BatMiner is a hybrid decision assistance system. Stages were integrated into RST-BatMiner. At first, we utilized the RST-based QUICK-REDUCT method to prune the dataset of extraneous characteristics. Then, the fitness function for each class was decreased, and BA was used to generate fuzzy rules. The rules generated by BA were then subjected to an Ada-Boosting technique, which is used to increase the precision of fuzzy rules. As a result, the BA encoding method was updated with a new operator to generate more thorough fuzzy rules. In addition, a unified fuzzy ruleset was built by identifying the rules associated with each class independently. Unfortunately, RST-BatMiner added extra time to the process of disease diagnosis.

3. Deep Neural Learning Classification

In order to enhance the efficiency and accuracy of early-stage diabetes diagnosis, the Logistic Fully Recurrent Deep Neural Learning Classification (LFRDNLC) method is developed. For the purpose of detecting the diabetic illness with the efficiency of classification, Longfei Han et al. [6] presented the ensemble learning strategy. Abril When it comes to efficiently classifying medical data, the hybrid associative classifier was put into place by Valeria Uriarte-Arcia et al. [8]. Yet, crucial improvements were not made to the accuracy of illness data classification. Hence, the suggested LFRDNLC method uses temporal sequences to get a comprehensive understanding of the dynamic changes present in patient medical data. Accurate predictions of the onset of diabetes are made possible by this deep learning. In contrast to other deep learning approaches, the proposed LFRDNLC method makes use of Fully Recurrent Neural Networks (FRNNs), in which every single neuron is linked to every single other neuron.

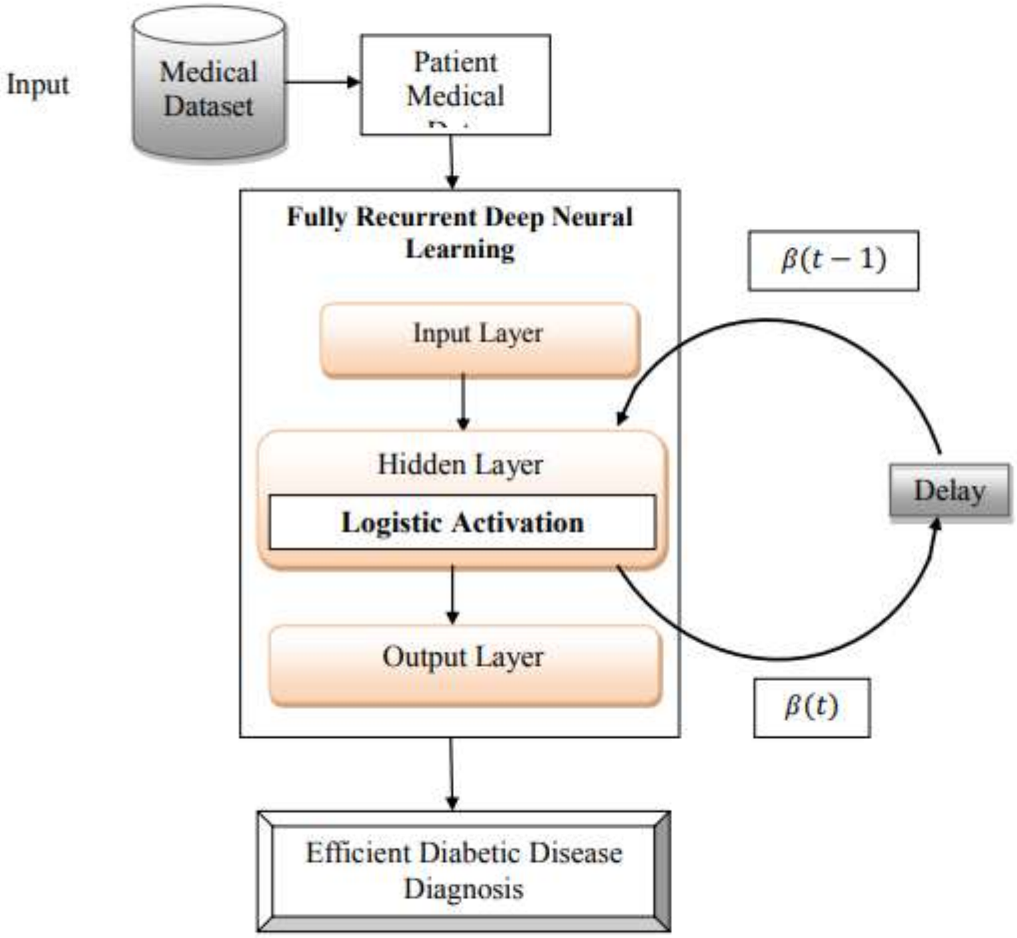

Figure.2: Architecture Diagram

In terms of dynamics, the FRNN is time-varying and nonlinear. So, the suggested LFRDNLC method may be used to accurately diagnose diabetes at an early stage by detecting the patient's symptoms at various time instances. Moreover, the FRNN is applied for feature classification (here, patient medical data) without regard to the feature's length. This aids in the automated determination of feature vector context. As a result, the suggested LFRDNLC method benefits greatly from the incorporation of FRNN in terms of improved classification performance. The suggested LFRDNLC methodology is a neuronal network (i.e. nodes). Each network node is directly connected to every other network node. The suggested LFRDNLC method aids in the identification of symptom changes in diabetes by having each neuron execute time-varying real-valued activation. The weight of each link is an actual number that may be changed. The suggested LFRDNLC method also includes three distinct varieties. The architecture of the Logistic Fully Recurrent Deep Neural Learning Classification method is shown in Figure.2. As seen in the diagram above. Second, the classification performances for diagnosing diabetes are improved by the suggested LFRDNLC method. The suggested LFRDNLC method first receives its input from the medical dataset (i.e., the diabetes dataset). Each patient's 50 characteristics across many time points in the medical dataset are sent to the input layer before being passed on to the hidden layer.

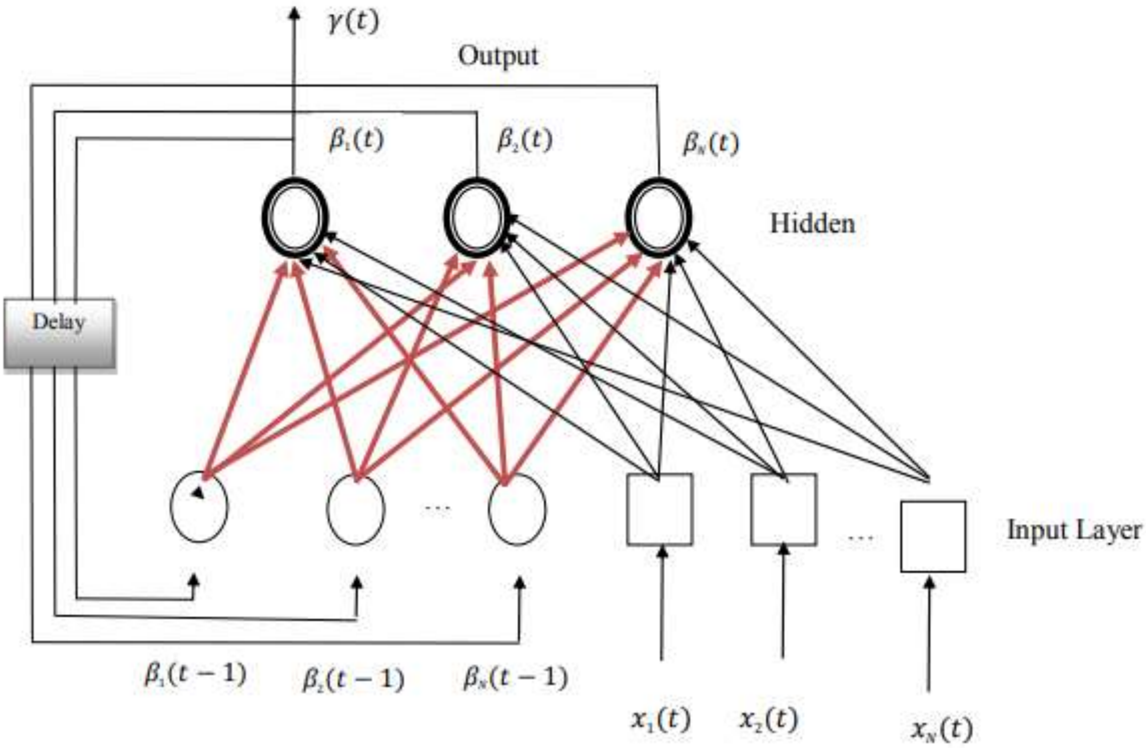

Figure.3: Type 2 Diabetes Prediction Using the LFRDNLC Method

In order to determine the association between the dependent variable (diabetes) and one or more independent variables (patient medical data) at a given time, the Sigmoidal Activation Function is inserted as a logistic regression function in the hidden layer. Thereafter, the inputs for assessing the patient's diabetes symptoms at various times are fed back into the network, together with the result of the hidden layer. This allows for quicker and more accurate diagnosis of the patient's condition. The hidden layer then sends the diagnostic result to the output layer, which processes it and returns the output results. Logistic Fully Recurrent Deep Neural Learning Classification is shown in its overall architecture in Figure.3. The proposed LFRDNLC method has a three-layer structure, as shown in figure.3. These layers are the input layer, the hidden layer, and the output layer. The input layer is completely coupled to the output layer through tunable weighted connections. The suggested LFRDNLC method also has unit-delay feedback connections that are fed back into the input layer. This aids in the accurate diagnosis of diabetes. Learning a mapping from one set of input sequences to another set of output sequences is how the aforementioned tasks are completed. To improve the accuracy of illness prediction categorization, [18] created an adaptive support vector machine. Nevertheless, patient data categorization did not achieve the optimal degree of efficiency. The suggested LFRDNLC method efficiently estimates the connection between diabetes illnesses and patient medical data by using the logistic activation function. This allows for a precise diagnosis to be made in a short period of time. The typical logistic sigmoid function's output is seen in Figure.4.

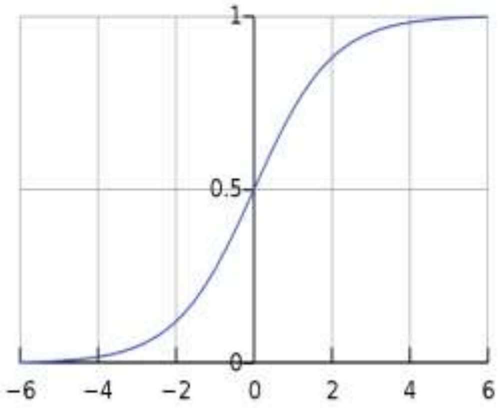

Figure.4: The Outcome of the Logical Activation Procedure

Figure.4 shows how the hidden layer makes use of the output of the logistic activation function. The logistic sigmoid function's output may take on values between 0 and 1 using this parameterization. Diabetic patients may be identified by their symptoms, which are taken into account by current diagnostic tools. The suggested LFRDNLC method is different from current approaches since it utilizes an unique logistic activation function to determine whether or not a patient is healthy. Each patient's medical records from various times are summarized using the present logistic sigmoid result and the previous hidden layer result.

4. Data Mining Techniques for Diabetic Disease Diagnosis

Knowledge may be uncovered via the identification and organization of massive amounts of data, and data mining is a crucial stage in this process. It has also shown great promise in the field of medical science by revealing previously undiscovered patterns in massive data sets. The medical field makes use of these patterns to provide more accurate diagnoses and improve patient care. Categorizing medical records is difficult because some information may be missing or the classifications may be wrong. When used for medical purposes, data mining aids in the creation of clinically informed judgments by medical professionals. Association rule mining is a data mining approach used to discover associations between patient medical records and diabetes symptoms. Classification is a data mining technique used to anticipate outcomes by building a model from historical information. For the purposes of organization, administration, and retrieval, medical records undergo a variety of categorization procedures. In order to increase classification performance, a hybrid platform must be implemented, since this cannot be done with a single algorithm. Effective medical diagnosis relies heavily on the ability to identify relevant information from medical databases. The primary goal of data mining is to extract meaningful patterns from databases and present them in an easily digestible fashion. Many association methods are used in the medical industry to derive rules from databases. Rules are developed after taking medical data bases into account. This set of rules is then compared to a set of standards. Information is often mined from medical databases using an association rule with a classifier. Better predictions were made thanks to this form of exploitation. When one group has a categorical characteristic and the other has two continuous-valued attributes, a contrast set is constructed to better illustrate the contrasts between the two groups. The flow process of creating association rules for diagnosing diabetes is shown in Figure.5.

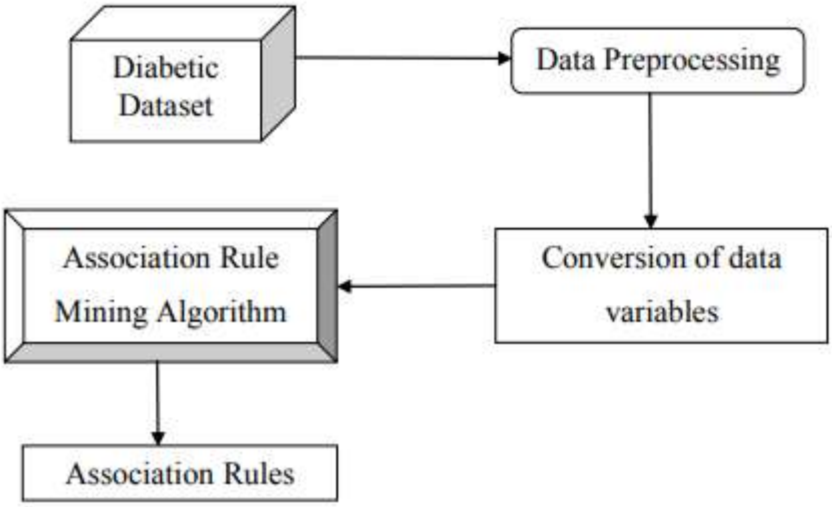

Figure.5: Flow Process

The quality of the data in the diabetes dataset may be improved with some pre-processing, as seen in the image. After data cleaning, categorical information is required for the categorization of broad relationships. The variables in the data are therefore transformed into a manageable set of categories. Finally, an association rule mining algorithm is used to extract reliable rules from healthcare datasets, along with a confidence score to aid in making an informed choice. Machine learning deals with the development of tools for data recognition. The primary need is for researchers to devise strategies for taking many patterns into account and drawing novel conclusions from existing information. The model is broken down into meaningful groups using classification and prediction operations to better understand and forecast their behavior. [19] created a machine-learning-based artificial-intelligence system for diabetic illness classification. Expectation maximization, principal component analysis, and support vector machine were used for the clustering, noise reduction, and classification processes. To deal with incremental conditions and incremental data recognition, respectively, incremental principal component analysis and incremental support vector machine were used. At the expense of computing time, prediction accuracy was improved, but classification accuracy was not. The diagnostic system uses the training data to apply rule recognition, and the test data to evaluate the model's performance. More precision-oriented approaches are needed for data processing and information recovery from medical datasets if we are to successfully contain significant challenges affecting human health. This is the primary motivation for using data mining techniques for dealing with diabetes categorization problems.

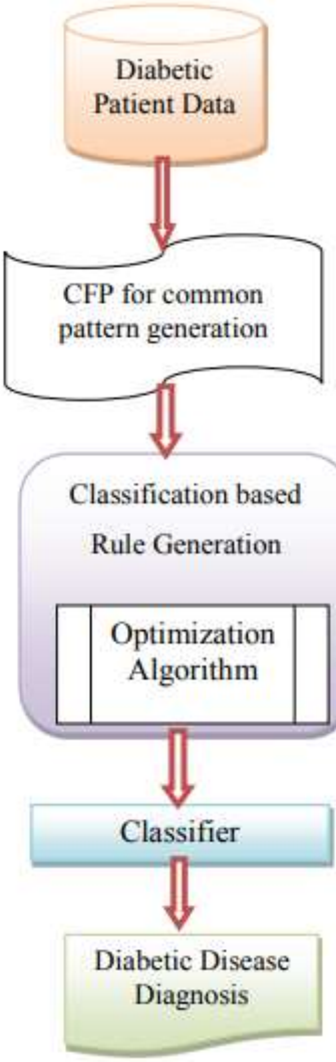

Figure.6: Diabetes Diagnose Classifications by Association

All individuals with diabetes fall into the Type 2 Diabetes group. Type 2 diabetes may have unexpected consequences, making it difficult to diagnose the disease based on symptoms alone. Data mining is modified for use in this context to perform the necessary analysis. Association rule mining was used to aid in decision-making. Using several minimum support values, complete frequent pattern (CFP) growth may be utilized to acquire frequent patterns. CFP built a MIS-tree (Multiple Item Support tree) and used it to produce shared patterns. CFP's overall performance is enhanced by using least minimal support and occasional child node pruning. The generated common patterns are optimized using an optimization approach. Finally, we employ a powerful classifier to extract insights from medical information collected from real patients. It also guided the rule mining process to limit the overlooked important rules generated by association rule mining methods. The Association Classification of Diabetes Diagnose is shown in Figure.6. Input data from diabetic patients is shown in figure.6. Next, redundant information is removed during preprocessing so that detection precision may be improved. Common Features Processing (CFP) is used to discover repeating structures in the data source. After that, an optimization technique was used to isolate the best CFP-derived association rules. Finally, a classifier is used to separate healthy data from ill data by using the created association rules. Association rule creation aids in identifying connections between the signs and symptoms of a disease and healthy patients because associative rules are more practical and realistic. Yet, there is still a lack of effective categorization methods for diagnosing diabetes.

5. Results And Discussion

Here, we compare the proposed Logistic Fully Recurrent Deep Neural Learning Classification (LFRDNLC) approach to others already on the market, such as the Ensemble Learning Method [10] and the Small-World Feed Forward Artificial Neural Network (SW-FFANN) [20]. The effectiveness of the suggested LFRDNLC method is determined using the numbers in the tables and graphs below, as well as the aforementioned metrics. The accuracy of patient categorization is analyzed for three distinct approaches using sample sizes ranging from twenty-five to two-hundred and fifty. In order to run the simulation, we compare the suggested LFRDNLC approach to others already in use, including the Ensemble Learning Method and the SW-FFANN. The accuracy of the classification is increased over time as more patients are added to the datasets used by the one suggested approach and the two existing methods. Figure.7 depicts a scatter plot of patient count vs test score classification precision.

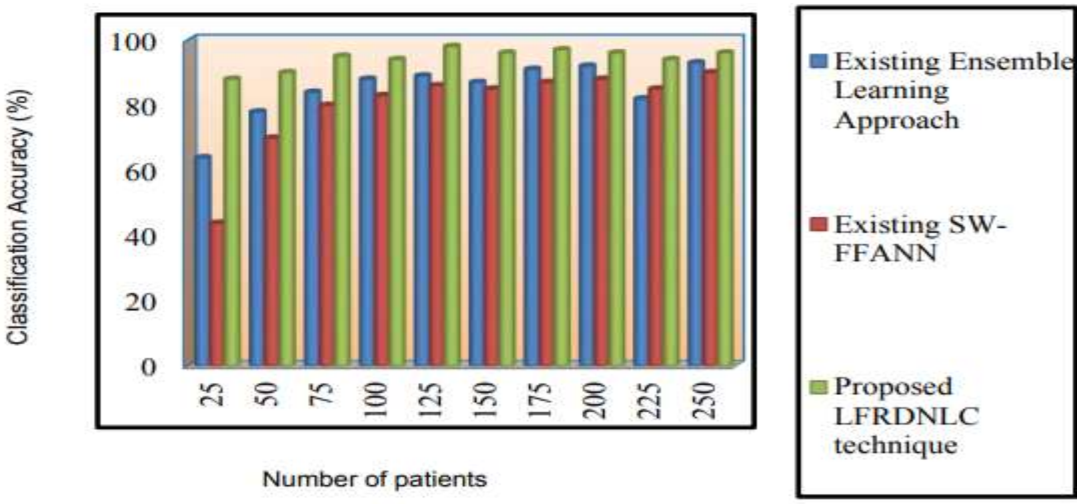

Figure.7: Measurement of Classification Accuracy

Classification accuracy versus patient count is illustrated in Figure.7. The simulation runs for 10 cycles, each time with a different patient count in the range of 25–250. By comparing the suggested LFRDNLC method with the Ensemble Learning Method and SW-FFANN, its efficacy in terms of classification accuracy is confirmed. When used to illness diagnosis, these three techniques improve classification performance.

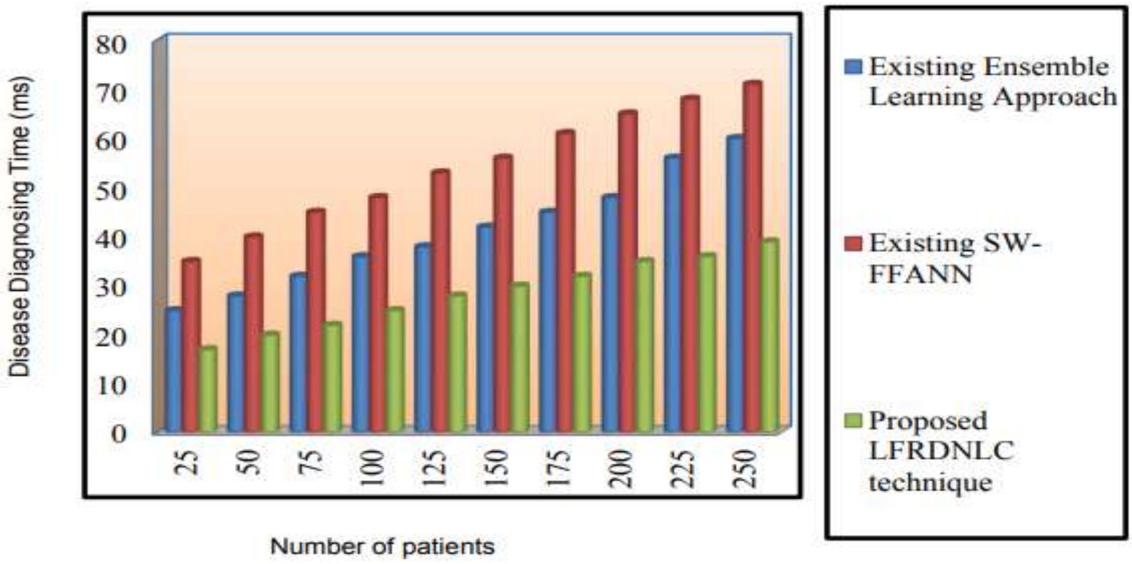

Figure. 8: Measurement of Disease Diagnosing Time

The suggested LFRDNLC approach improves classification accuracy more than any of the two competing techniques. This is because the logistic Sigmoidal function may be calculated during the diagnosis of diabetes using the suggested LFRDNLC method. Higher disease prediction performance may be achieved in disease diagnosis by using the logistic Sigmoidal function to assess the association between each patient's medical data and diabetic disease symptoms. The patient is deemed to be abnormal if the result of the output layer is 1, and to be normal otherwise. Furthermore, when compared to the current Ensemble Learning Method and the current SW-FFANN, the suggested LFRDNLC approach improves classification accuracy by up to 12% and 23%, respectively. Performance study of illness diagnosis time per patient for three approaches is shown in figure.8. Any number between twenty-five and two hundred and fifty patients is accepted as an input. The new LFRDNLC technique's performance is compared with that of the current techniques Ensemble Learning Approach and SW-FFANN to ensure its efficacy. As compared to current approaches, the suggested LFRDNLC technology speeds up the process of illness diagnosis. The suggested LFRDNLC method uses the calculation of the Logistic Sigmoidal Function to speed up the process of illness diagnosis. The link between patient data and indicators of diabetes illness is calculated using the logistic sigmoidal function. This paves the way for more precise early prediction of diabetes' existence or absence. As a result, less time is needed to confirm a diabetes diagnosis. As compared to the current Ensemble Learning Method and the current SW-FFANN, the suggested LFRDNLC method reduces the time required to diagnose an illness by 30% and 48%, respectively.

6. Conclusion

To better predict the onset of diabetes at an earlier stage and in less time, the Logistic Fully Recurrent Deep Neural Learning Classification (LFRDNLC) method is presented. Diabetic patients' fluctuating symptoms may be thoroughly learnt according to their various temporal sequences using the suggested LFRDNLC approach for disease diagnosis. This aids in making a correct diagnosis of the existence of illness. The suggested LFRDNLC method utilizes a three-layer structure, consisting of an input layer, a hidden layer, and an output layer, to accomplish the diagnosis of diabetes. The input layer receives the patient information from the data set and passes it on to the hidden layer. Second, to establish the connection between diabetes symptoms and a patient's medical data at a given moment, a Sigmoidal Activation Function-based logistic regression function is created in the hidden layer. The hidden layer's output is then combined with the inputs and used to train the network once more. This allows us to learn about the patient's diabetes symptoms over time. The data from the hidden layer is then sent to the output layer. When the output layer's result is 0, the suggested LFRDNLC method labels the patient as normal. If the output layer's result is 1, however, the suggested LFRDNLC method labels the patient as abnormal. Consequently, with the aid of the suggested LFRDNLC method, the onset of diabetes illnesses may be detected quickly and efficiently.

References

1. Aicha Boutorh and Ahmed Guessoum, "Complex diseases SNP selection and classification by hybrid Association Rule Mining and Artificial Neural Network-based Evolutionary Algorithms", Engineering Applications of Artificial Intelligence, Elsevier, Volume 51, May 2016, Pages 58-70.
2. Biao Jie, Mingxia Liu, Daoqiang Zhang, Dinggang Shen, "Sub-Network Kernels for Measuring Similarity of Brain Connectivity Networks in Disease Diagnosis", IEEE Transactions on Image Processing, Volume 27, Issue No.5, May 2018, Pages 2340 – 2353.
3. Bo Zheng, Hong-Zhong Huanga, Wei Guo, Yan-Feng Li and Jinhua Mi, "Fault diagnosis method based on supervised particle swarm optimization classification algorithm", Intelligent Data Analysis, Volume 22, 2018, Pages 191–210.
4. Emre Gurbuz, and Erdal Kılıc, "A new adaptive support vector machine for diagnosis of diseases", Expert Systems, Wiley Online library, Volume 31, Issue No.5, November 2014, Pages 389–397.
5. Fei Ye, "Evolving the SVM model based on a hybrid method using swarm optimization techniques in combination with a genetic algorithm for medical diagnosis", Multimedia Tools and Applications, Springer, Volume 77, Issue No.3, February 2018, Pages 3889–3918.
6. G. Thippa Reddy and Neelu Khare, "FFBAT- Optimized Rule Based Fuzzy Logic Classifier for Diabetes", International Journal of Engineering Research in Africa, Volume 24, June 2016, Pages 137-152
7. Hsin-Yi Tsao, Pei-Ying Chan and Emily Chia-Yu Su, "Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms", BMC Bioinformatics, Volume 19, 2018, Pages 195–205
8. Jyotsna Bansal, Divakar Singh and Anju Singh, "An Efficient Medical Data Classification based on Ant Colony Optimization", International Journal of Computer Applications, Volume 87, Issue No.10, February 2014, Pages 14-19.
9. Kemal Polat, SalihGüneş, AhmetArslan, "A cascade learning system for classification of diabetes disease: Generalized Discriminant Analysis and Least Square Support Vector Machine", Expert Systems with Applications, Elsevier, Volume 34, Issue No.1, January 2008, Pages 482-487.
10. Longfei Han, Senlin Luo, Jianmin Yu, Limin Pan, Songjing Chen, "Rule Extraction From Support Vector Machines Using Ensemble Learning Approach: An Application for Diagnosis of Diabetes", IEEE Journal of Biomedical and Health Informatics, Volume 19, Issue No.2, March 2015, Pages 728 – 734.
11. M Nandhini and S N Sivanandam, "An improved predictive association rule based classifier using gain ratio and T-test for health care data diagnosis", Sadhana, Springer, Volume 40, Issue No.6, September 2015, Pages 1683–1699.
12. Madhu Sudana Rao Nalluri, Kannan K, Manisha M and Diptendu Sinha Roy, "Hybrid Disease Diagnosis Using Multiobjective Optimization with Evolutionary Parameter Optimization", Journal of Healthcare Engineering, Hindawi, 2017, Pages 1-27.

13. Nahla Barakat, Andrew P. Bradley and Mohamed Nabil H. Barakat, "Intelligible Support Vector Machines for Diagnosis of Diabetes Mellitus", IEEE Transactions on Information Technology in Biomedicine, Volume 14, Issue No.4, July 2010, Pages 1114 – 1120.
14. Peng Cao, Fulong Ren, Chao Wan, Jinzhu Yang and Osmar Zaiane, "Efficient multi-kernel multi-instance learning using weakly supervised and imbalanced data for diabetic retinopathy diagnosis", Computerized Medical Imaging and Graphics, Elsevier, Volume 69, November 2018, Pages 112-124.
15. Ramalingaswamy Cheruku, Damodar Reddy Edla, Venkatanareshbabu Kuppili, Ramesh Dharavath, "RST-BatMiner: A Fuzzy Rule Miner Integrating Rough Set Feature Selection and Bat Optimization for Detection of Diabetes Disease", Applied Soft Computing, Elsevier, Volume 67, June 2018, Pages 764-780.
16. Salim Amour Diwani and Anael Sam, "Diabetes Forecasting Using Supervised Learning Techniques", ACSIJ Advances in Computer Science: an International Journal, Volume 3, Issue No.11, September 2014, Pages 10-18.
17. Seyed Abbas Mahmoodi, Kamal Mirzaie and Seyed Mostafa Mahmoudi, "A new algorithm to extract hidden rules of gastric cancer data based on ontology", Springer Plus, Volume 5, Issue No.312, 2016, Pages 1-21.
18. Sujata Joshi, Mydhili K. Nair, "Prediction of Heart Disease Using Classification Based Data Mining Techniques", Computational Intelligence in Data Mining, Springer, Volume 2, December 2014, Pages 503-511.
19. Xiao Liu, Xiaoli Wang, Qiang Su, Mo Zhang, Yanhong Zhu, Qiugen Wang, and Qian Wang, "A Hybrid Classification System for Heart Disease Diagnosis Based on the RFRS Method", Hindawi Publishing Corporation, Computational and Mathematical Methods in Medicine, Volume 2017, Pages 1-11.
20. Anuradha, Singh, Akansha, Gupta and Gaurav, "ANT_FDCSM: A novel fuzzy rule miner derived from ant colony meta-heuristic for diagnosis of diabetic patients", Journal of Intelligent & Fuzzy Systems, IOS Press, 2018, Pages 1-14.