Performance Evaluation of Naive Bayes and Back Propagation Neural Network classifiers in gestational Diabetes Mellitus Classification

Geetha.V.R [1], Dr.Jayaveeran.N[2]

[1] Assistant Professor, Department of Computer Science, A.V.C College (Autonomous), Mannampandal,
Tamil Nadu, India

[2] Former HOD of Department of Computer Science, KhadirMohideen College, Adhirampattinam,
Tamil Nadu, India

(Affiliated to Bharathidasan University, Tiruchirappalli- 620 024, Tamil Nadu, India)

ABSTRACT

The classification of Gestational Diabetes plays significant role in healthcare domain. Gestational Diabetes Mellitus (GDM) is a condition in which a hormone made by the placenta prevents the body from using insulin effectively. Glucose builds up in the blood instead of being absorbed by the cells.GDM is getting more common as overweight as well as next medical disorders are growing. It needs diligent and knowledgeable medical supervision, because the treatment provided throughout pregnancy impacts not only obstetric safety but also the likelihood of growing type 2 disease many years into the further, in both the infant and the mother. In this paper, precision of GDM classification is determined by comparing Back Propagation Neural Network (BPNN) and Enhanced Naïve Bayes classifiers. BPNN classifier is a type of ANN classifier that is developed for classification of GDM disease with accuracy estimation. The Enhanced Naïve Bayes classifier is the ML classification technique that is developed. The classification assessment methods Confusion Matrix (CM) and Balanced Error Rate(BER) are used to compare both the classification methodologies' efficiency and accuracy. The experimental results show that proposed Enhanced Navie Bayes methodology is having higher degree of efficiency and accuracy.

Key words : Gestational Diabetes Mellitus (GDM), Artificial Neural Network (ANN), Machine Learning (ML), Naïve Bayes classifier, Confusion Matrix (CM), Balanced Error Rate(BER).

INTRODUCTION

It is well known that GDM incidence differs based on the surveyed group and the clinical parameters used. GDM is considered to induce disease that is caused by pregnancy. Pregnancy lessens the release of glucose. In medium or late pregnancy this often emerges in illumination. The high insulin rate is likely to be transferred via uterus to the infant, which can limit infant development. With a stable child, GDM must be managed as soon as it is identified. The classification system that constructs the research should anticipate the development of GDM based on early maternity variables. The information source used may also trigger major variations in GDM incidence monitoring but is regarded less regularly. GDM incidence differs greatly based on the nature of the information being accessed. Enhanced patient coordination may help female self-manage GDM [1]. GDM is a significant modifiable risk factor for negative effects in birth and is correlated with an elevated chance in the subsequent life in acquiring type 2 diseases. This susceptibility underlying chronic inflammation will be further compounded by the substantial drop in glucose responsiveness during regular gestation which, in conjunction with alpha-cell malfunction, results in the production of GDM [2]. GDM a temporary type of pregnancy-induced diabetes-has more significant short- and long-term health effects for both mother and her fetal. A GDM disorder has far-reaching implications that are specific to a large community of people. Blood sugar reduction cut-offs cause the possibility of physical damage to mother and baby. Consideration of the dangers and rewards from a behavioral and a biochemical standpoint can also be beneficial. This will reduce additional pressure on a community increasingly insecure. Postpartum hemorrhage (described as a neonate measuring more than 4 kg) is the most commonly recorded obstetric effect of GDM, and may raise the likelihood of cesarean section and joint postpartum hemorrhage. There are several possible larger-term effects for the parent involving an elevated likelihood of type 2 disease during birth and/or entire life [3]. GDM is a top clinical disorder in which people face severe short- and long-term effects for pregnancy morbidity throughout infancy. The probability of harmful pregnancy consequences improve with GDM. That entails improving prenatal care and treatment

approaches for mothers with GDM. Although it is possible to enhance the adverse maternal consequences of mother with GDM by adequate prenatal treatment and constructive improvements in lifestyle, the likelihood of negative outcomes improve drastically as a consequence of elevated maternal glucose rates in the second or third pregnancy, also beyond thresholds commonly deemed common for pregnancy. Though there are effects of GDM on harmful reproductive results in developing nations with numerous conflicts, there is minimal information on the impact of GDM [4].Despite the absence of globally recognized GDM- related diet recommendations, its objective is to give adequate nutrition to support reproductive and fetal wellbeing while at the same time to help and meet glycogenic targets and it relies on a level of creative supplementation. Biochemical therapy is undertaken where that is not enough [5]. Artificial Neural Network (ANN) is used to classify pre-diabetes and phase 2 diseases (T2D). A lattice neural network framework equipped by the Levenberg–Marquardt (LM) methodology and a deterministic neural network framework detects the diabetic condition by careful analysis of the insulin results[6].On the Canadian community, an efficient statistical model with high precision and specificity is built and tested and is more accurate and effective to adapt to Canadian clinicians than current models produced from US or other societies. The most powerful determinants in these simulations were fasted blood sugar, human body mass index, high degree of density lipoprotein and cholesterol [7].The planned research develops on early diagnosis of GDM for mother who is second time pregnant without a referral to clinic. The need of Artificial Neural Network technologies for disease prediction is showing improved efficiency in ethical decision-making. ANN is commonly used for detection and treatment of different fields of hunger such as efficient patient decision taking, image processing, and so on. Artificial neural networks are used in metabolic research to evaluate blood samples, which monitor glucose variable in disease [8]. Categorization questions in the deep-learning research have a long tradition. The Enhanced Naïve Bayes systems are one of the easiest, yet most reliably excellently-performing groups of algorithms. Hierarchical enhanced Bayes systems expand the versatility of simulation of Naïve Bayes systems by adding conditional parameters to reduce some of the claims of freedom in these systems [9].Naïve Bayes learners in confined areas are commonly utilized, reliable, and accurate controlled forms of learning for labeled databases. It has been shown that learners from Naïve Bayes deliver good efficiency relative to other algorithms for artificial intelligence. A collection of tests are performed on 17 UCI comparison databases to equate the performance of the suggested learner with that of other approaches using a flexible subjective presumption of freedom. The ultimate findings often show the supremacy of approaches for attribute-weighing over those that seek to assess the system architecture [10].

Related Works

Yashi Srivastava et. al [11] has investigated Diabetes is a disease that gradually destroys the structures of natural protection and internal organs; it is a quiet and deliberate destruction. It is a situation once due to heavy blood sugar levels immune system fails to exploit the energy generated by food. Diabetes is one of the possible causes of kidney failure, cardiac arrest, impairment and surgical removal of the limbs.Amita Mahajan et. al [12] has performed a good quality comprehensive studies which reveal no substantial differences in small-carbohydrate or limited-calorie diets. Ricardo Augusto Leoni De Sousa [13] has proposed A type of sugar sensitivity can be described as GDM. Objective evidence for late physiological and behavioural implications is appearing for children born from GDM. More latest and related evidence is pointing to adverse effects of GDM on the children's attitude and brain function.Jasmine F. Plows et. al [14] have investigated GDM is a serious complication of pregnancy that has short- and long-term health consequences for mom and infant alike.Santosh Kumar et. al [15] have investigated data processing methods commonly used to collect insightful details from repositories of scientific records. Categorization in machine learning is a supervised learning that can be used to build techniques representing essential groups of information, where category element is included in the clustering algorithm creation. Naïve Bayes is a very basic, common, highly effective, and successful machine learning technique. In this work, it is suggested to integrate creative method Naïve Bayes with clustering algorithm for successful identification.Md. Maniruzzamanet. al [16] have investigated Diabetes was treated as a medical condition marked by elevated blood sugar. It can contribute to multiple complex illnesses such as diabetes, kidney disease, cardiac arrest, etc. The major purpose of this study is to build a program focused on machine learning (ML) for forecasting gestational diabetics. Cross-validation (CV) protocol is recognized as code splitting. This is used specifically to split the specified database into two subgroups, such as: I training set and (ii) approval set. There are several CV algorithms that are used to segment the database to through heterogeneity. The tenfold CV procedure is widely used in ML and analytics, where the database is split into ten equivalent sections while the nine pieces are used as an ML-based device clustering algorithm and the remaining portion is used as an approval / test collection.

AbidSarwaret. al [17] havedescribed Artificial Intelligence (AI) which is a core technology in several sectors is now becoming increasingly prevalent in clinical diagnosis, and has been commonly utilized in the treatment of cancers, diseases, tuberculosis, respiratory diseases, etc. Authors defined ten variables that play a significant role in disease and provided a rich set of dataset that acted as the foundation of

the architectures for forecasting. Holding this training data in mind, researchers introduced three Bayes, Artificial Neural Networks (ANN), and K-nearest Neighbors (KNN) protocols, and established correlation systems. To measure the efficacy, forecasting method tests were correlated with the subjects' real clinical diagnosis. To carry out the work described in this work, researchers constructed a sample collection comprising sample on 500 individuals selected randomly from various parts of society so that information diversity could be assured. In the data collection the average and minimum age of respondents is 5 and 78 years, accordingly. The data collection is composed of ten physiological variables that play a significant role in diabetic realization. After thorough consideration with interested sector experts, these criteria were selected. The variables were given separate variables so as to ensure accuracy in the system. Significant consideration was taken to insure that the records had accurate attributes to guarantee consistency and to prevent mistakes.

K-nearest neighbor (KNN) is an instance-dependent grouping methodology that categorizes the test items based on the amount of learning instances nearest to them. It is a nonparametric methodology which means it makes no inference about the fundamental representation of the results. K is a typically very low positive amount. More challenging is to differentiate between the different types, the increasing the importance of K. Specific heuristic methods, including cross-validation, are used to make a successful choice of K.Harry Zhang et. al [18] have proposed Naive Bayes which operates remarkably well in identification but its estimate of likelihood is weak. A classification dependent on class variables is therefore required in many implementations. Our studies reveal that naive Bayes exceeds C4.4, the most state-of-the-art implementation in decision tree rankings. We are discussing two instances of issues that were used in identification research of the results of naive Bayes. Particularly, Bayes succeeds well in rating on them, even when it is not in ranks.Peter A. Flachet. al [19] have investigated1BC and 1BC2, two programs conducting the naive Bayesian analysis of organized entities. 1BC's solution is to project the entities into characteristics of the first level. These attributes are constructed from the entity using functional predicates relating to similar objects (e.g., atoms inside molecules), and characteristics belonging to the person or to one or more of his similar items (e.g., a connection among two molecules).Yuqian Jiang et. al [20] has proposed that Naive Bayes classifier is commonly used in text identification projects, and can work fairly well, it is also called a benchmark. But previous work indicates that the distorted composition of the set of learning may produce poor results in classifying text. The author offers a different way of coping with the case.

OBJECTIVE

GDM is a metabolic condition which is rapidly widespread, marked by the failure of the organism to metabolize insulin. The goal of this research is to identify an efficient classification model with high accuracy and sensitivity to help classify patients at risk of developing GDM based on clinical demographic information and the findings of the testing throughout their trips to medical centers. This model forecasts diabetes patients using certain widely used laboratory findings with satisfying tolerance. The GDM classification is carried out by two classifiers named Back Propagation Neural Network (BPNN) type of Artificial Neural Network and enhanced Naïve Bayes classifier part of Machine Learning (ML). The Confusion Matrix (CM) and Balanced Error Rate(BER) are used to test the classification efficiency and accuracy in percentage level.

Proposed Methodology

A machine learning algorithm is introduced to a GDM prediction or identification function. These models either sought to categorize clinical patients into glucose and non-glucose, or they predicted the intensity of blood increase in patients. Most health experts have discovered there is a good relationship between the effects of clinicians with certain chronic deficiency and the levels of insulin. This research presents a forecast-description model of GDM classification in the form of two algorithms to validate the association. The first sub-module uses Machine Learning (ML) and second uses Artificial Neural Network (ANN) for classifying case categories and predicting patient GDM levels.

System Model

The Artificial Neuron Network (ANN) widely known as the Neural Network is a mathematical formalism focused on brain neuron neural network configuration and features. It is like an automated brain and nervous system that collects processed and transmits knowledge from an engineering viewpoint.

Generally, a neural network comprises of 3 distinct layers:-
- Data layer (Over this level all information is fed into the system)
- Hidden layers (More than 1 secret layers may be used to process information obtained from the layers of data)
- Output layer (Information is rendered accessible at the output layer during refining)

ANNs are simulated dynamic mechanisms which are influenced by the human mind's working systems. They are structures which can change their inner structure in response to an analytical feature. They are especially useful for tackling

dynamical style challenges, becoming able to recreate the fuzzy principles which regulate the appropriate solution to this issue. The ANN's foundation components are the branches and the relations, also named processing elements (PE). Each neuron has its own origin, from which it collects contact from other entities and/or from the world and its own performance, through which it interacts with other clusters.
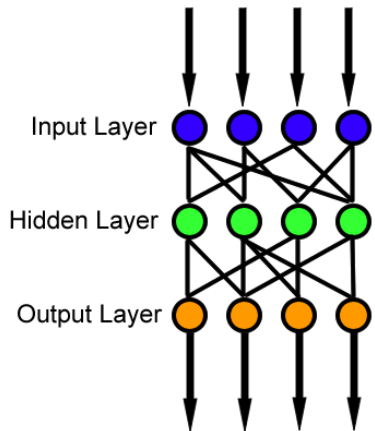

Figure 1: Neural Network Layers

The characteristics of ANNs are strong in object recognition-like capabilities that are required for information processing and prediction taking, and are versatile optimization algorithms with the capacity to categorize and take decisions from broad and quite fuzzy input data. ANNs are especially specialized in designing processes for resolving nonlinear issues. In mathematical words, we can assume that when the structure describing it is static, a mechanism isn't complicated, that is, when these two parameters apply:

$$F(AX)=AX(F) \tag{1}$$
$$F(A1+A2)=F(A1)+F(A2) \tag{2}$$

A Complicated, dynamical structure infringes one or more of those words. In general, the more dynamical the feature, the more useful it is to use an ANN to try to explain the law, R, which controls the action within the black box. If we take a Euclidean map in which angle x reflects the sum of cost that a state gets and angle y calculates the degree of satisfaction the person gets as a consequence, then y axis reflection gets output data through processing by hidden layer of ANN.

ANN hidden rules

ANNs are systems for information processing and do not obey specific data processing principles but which use information they collect to uncover the laws controlling them. It allows ANNs especially helpful when it comes to addressing a question with which we have the information associated, but may not know how such information is connected. That

implies that if they obtain some data in one step, ANNs can rely on particular laws; but if they obtain fresh and specific results later, ANNs can change their laws appropriately, combining the old information with the latest, and they will do so with no explicit guidance. Sustained information upgrading within their supervision produces a fluid system, whose laws are continuously modified by the ANNs as topic grows over period. The ANN only handles this transition from such an earlier classification to a later, better, and more nuanced one, utilizing the existing instances as evidence to know about the current definition.

Back Propagation Neural Network Classification for GDM

The neural network protocol Back propagation (BP) is a multi-layer feed forward network equipped as per the method of loss back propagation which is one of the many commonly utilized neural network systems. BP system could be used to know and preserve a number of input-output routing relationships, with no need to reveal the mathematical formula defining such routing relationships in preparation. The training guideline is to follow the fastest decline approach under which the back distribution is used to control the system weight vector and limit vector to obtain the square's minimal error number. BP methodology is simply a form of tracking information. It uses mean squared error and differential regression approaches to allow the adjustment to node relation weight. The adjustment to network link weight is intended to obtain the lowest square error number. In this method, the system relation variable is given a little weight first, and then a testing test is chosen to measure the loss coefficient proportional to this test.
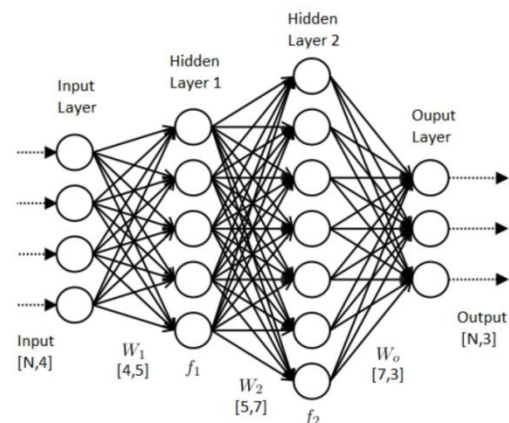

Figure 2: Back Propagation Neural Network Architecture

The Back Propagation Neural Network Activation Function

The neuron input is calculated as the weighted variable sum written by equation 3.

$$n = \sum_{i=1}^{r} AO_{Bi}CW_i \qquad (3)$$

In Figure 3, F is the commencement utility, which has a sigmoid function. Sigmoid commencement occupation is shown in Figure 3.

The flexibility of the sigmoid feature relative supports its importance and is used in learning architectures as an initialization feature. The performance of the neuronal is given by equation 3 and equation 4, with a sigmoid stimulation feature.

Output = F(n)   (4)
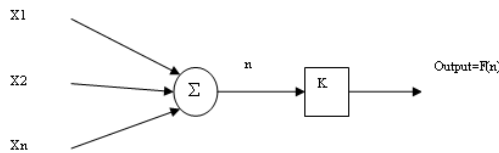
function(K) =   (5)

Figure 3 Structure of Artificial Neuron

The derivation of the sigmoid occupation can be calculated as follows Equation 6:

$$\frac{\partial F(K)}{\partial n} = output * (1 - output) = FKA) * (1 - K(A)) \qquad (6)$$
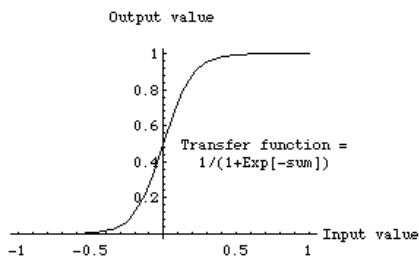

Figure 4 Sigmoid Activation Function

Machine Learning (Ml) Classifier

When a methodology trains from instance information and related goal outcomes which may comprise of quantitative variables or object identifiers, such as categories or marks, in attempt to later determine the appropriate answer when presented with new instances falls under the supervised training classification. However, the method is similar to human training below a teacher's guidance. For the participant to memorize, the instructor gives clear definitions, and the student instead draws general principles from such particular examples. An algorithm trains from simple examples without any accompanying answer, allowing the machine to evaluate the trends of information by itself. This type of methodology tries to redesign the information into anything else, such as innovative features that can reflect a group, or a new series of highly correlated principles. They are very helpful in offering insights into the importance of data and new useful contributions to supervised machine learning methodology for classification. As a kind of thinking, it parallels the methods employed by humans to work out that particular structures or things are of the same sort, for example by analyzing the level of resemblance between entities.

Enhanced Naïve Bayes Classifier for GDM

The NaïveBayes (NB) algorithm is an efficient and descriptive classification algorithm and is especially useful in large datasets. It is used in both machine learning and modern science (in particular, diabetes diagnosis). It is a deterministic clustering algorithm centered on the principle of NaïveBayes, with the clear autonomous expectation among the features. The inclusion of particular characteristics in a group is believed to be irrelevant to any other characteristics.Naive Bayes is a methodology of identification that describes all characteristics as separate and unrelated to one another. It determines that the position of a particular character in a category does not influence the position of another attribute. Since it is focused on maximum likelihood it is called efficient methodology used for the function of identification. With misbalancing issues and lacking qualities it functions well for the results. Enhanced Naive Bayes is a clustering algorithm for machine learning that uses the Bayes Principle.

The key aim of this research is to show the most successful identification using machine learning methodology based on enhanced NaïveBayes classification which is required to forecast early stage patients with GDM. There has been ample theoretical focus on intrusive, digital diabetes exploration. Therefore, it all relies on what characteristics were derived and on which variety of discriminator to get the optimum result was extended. Accordingly, Learning Complexity has been examined as such data collection variables may be used to identify different predictive health conditions. The alliance of disparate classification protocols tested on the different layers should establish a well-organized framework for diabetes exploration together with controlling the risk measurements and clinical practitioners' management strategies. This approach for classifying enhanced NaïveBayes comprises of two key sections, first of all how precision is achieved using different classification systems and secondary of all is system verification. There are various machine learning frameworks accessible that are useful in evaluating the unrecognized trends for risk factor analysis in infections such as diabetes. Furthermore, it is found that the analysis of traditional approaches owing to a large data aspect is not up to the degree of approval in recognition. ML is a model focused on multilayer feed-forward perceptrons that also promotes the characteristics of ANN and is equipped utilizing back-propagation with probabilistic differential decay. The system is a series of four node and neuron imitating layers, guided in uni-direction (one-way

association). Growing node is linked to another network in a possible way and includes two secret layers where each network uses local data to train a copy of global system specifications. It often requires several threads to analyze the application and add the average to connect across the entire network to the system entry. The training system uses probabilistic regression downward learning utilizing neuron from the secret layer of enhanced NaïveBayes that allow for more advancement functionality such as tanh, rectifier, and peak out sensitivity, training and, and annealing.

Enhanced Naïve Bayes Formulation

$$B(Ca, br_1, br_2, ..., br_M) = B(Ca) \prod_{i=1}^{M} B(ar_i \mid bC)$$

Then, the joint probability for the rating patterns of user y, i.e.

$\{aR_y(xx_1), bR_y(xx_2), ..., xR_y(xx_M)\}$, can be expanded as:

$$P(R_y(x_1), R_y(x_2), ..., R_y(x_M)) = \sum_C P(C) \prod_{i \in X(y)} P(R_y(x_i) \mid C)$$

This approach will initially select a type of function 'C' through the P(C) continuum as seen by Formula (8) and instead rate all entities utilizing the specified class 'C'. This model assumes, in other terms, a user group applied to score of all the symbols and thus eliminates the case where a user is of various classes and categories are assigned to the scores of various items.

RESULTS AND DISCUSSION

In this section, GDM classification using BPNN and enhanced NaïveBayes classification techniques are compared in terms of mean square error, efficiency and accuracy.
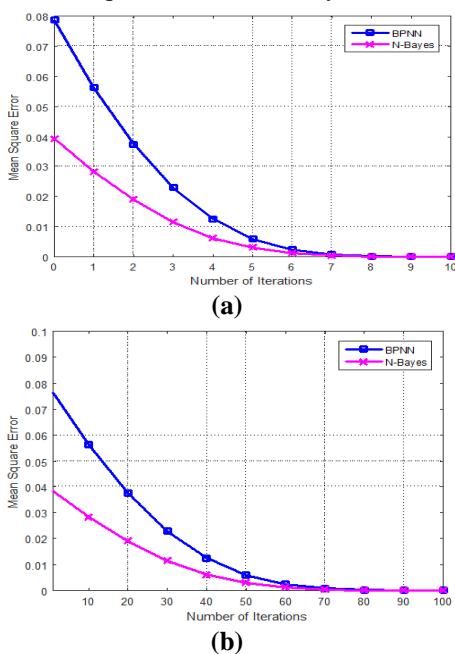

(a)

(b)

Figure 1 (a) Mean Square Error (MSE) Estimation for Iterations 10, (b) Mean Square Error (MSE) Estimation for Iterations 100.
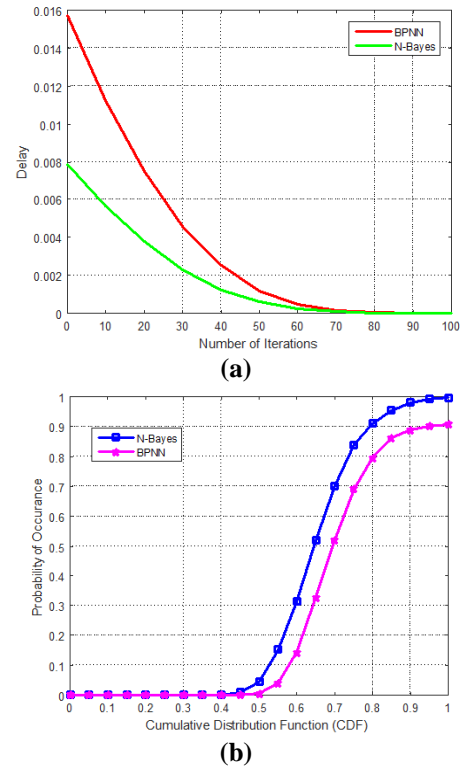

(a)

(b)

Figure 2 (a) Delay Estimation for BPNN and N-Bayes Classifier, (b) Probability of Occurance of BPNN and Enhanced Naïve Bayesclassifier
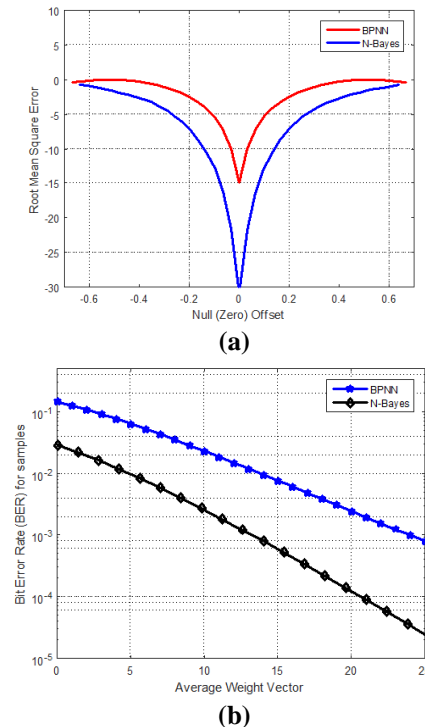

(a)

(b)

Figure 3: (a) Null Offset Correction for BPNN and N-Bayes Classification, (b) BER Minimization for Training Samples using BPNN and Enhanced Naïve Bayesclassifier

Figure 1 (a) shows the Mean Square Error (MSE) estimation for number of epochs or iterations of 10 using BPNN and enhanced NaïveBayes classifier. The MSE for enhanced N-Bayes classifier is highly reduced when compared to BPNN MSE error minimization. Figure 1 (b) shows MSE minimization for number of epochs or iterations of 100.Figure 2 (a) shows delay (Latency) computation for BPNN and enhanced N-Bayes classifier techniques. The enhanced N-Bayes classification delay is highly reduced while compared to BPNN as shown in figure 2 (a). Probability of occurrence is represented for BPNN and Enhanced N-Bayes classifier in figure 2 (b).Figure 3 (a) shows offset correction to zero for BPNN and N-Bayes classification to make the network sigmoid function activation for training and testing samples of GDM. The BER minimization for training samples using BPNN and enhanced N-Bayes classification is shown in figure 3 (b). The BER is minimized for enhanced N-Bayes classification than BPNN.

TABLE 1: ACCURACY TESTING USING CONFUSION MATRIX (CM)

S.No	Testing	Description	CM Formula
1	Precision (P)	Classifier correctness/accuracy is measured by Precision	P = TP / (TP+ FP)
2	Accuracy (A)	Accuracy determines the accuracy of the algorithm in predicting instances.	A=(TP+ TN) / (Total no of samples)
3	Precision (P)	Classifier correctness/accuracy is measured by Precision.	P = TP / (TP+ FP)

TABLE 2 PERFORMANCE ANALYSIS USING CM

Classification	Algorithms	Precision In %	Accuracy In %	Precision In %
ANN	BPNN	83	89	91
ML	Enhanced N Bayes	93	97	96

CONCLUSION

The classification of Gestational Diabetes (GD) is carried out using ANN and ML techniques. The BPNN algorithm is used as ANN model and Enhanced N-Bayes classification algorithm is used as ML Techniques. The graphical representation and tabulation results using Confusion Matrix (CM) show that the Enhanced N-Bayes classification algorithm performance is better than BPNN algorithm. The efficiency and accuracy of Enhanced N-Bayes classifier is better than BPNN.

REFERENCES

1. Robyn L. Lawrence, Clare R. Wall and Frank H. Bloomfield, "Prevalence of gestational diabetes according to commonly used data sources: an observational study", https://doi.org/10.1186/s12884-019-2521-2.
2. KatrienBenhalima , Pau Van Crombrugge, CarolienMoyson, Johan Verhaeghe, SofieVandeginste, Hilde Verlaenen, Chris Vercammen, Toon Maes, ElsDufraimont, Christophe De Block, Yves Jacquemyn, Farah Mekahli, Katrien De Clippel, Annick Van Den Bruel, Anne Loccufier, AnnouschkaLaenen, Caro Minschart, Roland Devlieger and Chantal Mathieu, "Characteristics and pregnancy outcomes across gestational diabetes mellitus subtypes based on insulin resistance", https://doi.org/10.1007/s00125-019-4961-7
3. Vidhya, K. "The role of serum calcium and zinc in pregnancy induced hypertension." TJPRC: International Journal of Pharmacology and Physiology 2.1 (2016): 1-14.
4. Louise Craig, Rebecca Sims, Paul Glasziou and Rae Thomas, "Women's experiences of a diagnosis of gestational diabetes mellitus: a systematic review", https://doi.org/10.1186/s12884-020-2745-1.
5. NAVJOTKAUR, HARI SINGH, and ANAND NAYYAR. "FUZZY LOGICS BASED EXPERT SYSTEM FOR DIAGNOSING DIABETES: A REVIEW." Development (IJECIERD) 3.3 (2013): 25-34.
6. AchenefAsmamawMuche, Oladapo O. Olayemi and YigzawKebedeGete, "Effects of gestational diabetes mellitus on risk of adverse maternal outcomes: a prospective cohort study in Northwest Ethiopia", https://doi.org/10.1186/s12884-020-2759-8.
7. Ines Morais Rodrigues, Ana Figueiredo, Naiegal Pereira, NjilaAmaral, SóniaPratas,CristinaValadas, Elsa Dias and Carlos Veríssimo. "Metformin as a Safe Option to Insulin in Gestational Diabetes Mellitus: a Retrospective Study", https://doi.org/10.1007/s42399-020-00229-9.
8. Choudhury, Purobi, and CHANDAN KUMAR DAS. "Periodontal disease and pregnancy outcome: a correlative study." Intern. J. Dental Res. Develop 7.2 (2017): 1-6.
9. DijanaSejdinovi, LejlaGurbeta, AlmirBadnjevi, Maja Malenica, Tanja Duji, Adlijaauševi, Tamer Bego, LejlaDivovi and Mehmedovi, "Classification Of Prediabetes And Type 2 Diabetes Using Artificial Neural Network", DOI: 10.1007/978-981-10-4166-2_103.

10. Hang Lai, Huaxiong Huang, Karim Keshavjee, Aziz Guergachi and Xin Gao, "Predictive models for diabetes mellitus using machine learning techniques", https://doi.org/10.1186/s12902-019-0436-6.

11. EZEBUENYI, MICHAEL, et al. "Evaluation of Selected Medicinal Herbs for Antidiabetic Activity via Alpha-glucosidase Inhibition." International journal of general medicine and Pharmacy vol 6 (2017): 59-64.

12. Priya Shirley Muller, M.Nirmala, "Diagnosis of Gestational Diabetes Mellitus using Radial Basis Function", IEEE 2016.

13. Helge Langseth and Thomas D. Nielsen, "Classification using Hierarchical Naıve Bayes models", DOI 10.1007/s10994-006-6136-2

14. Zhong-Liang Xiang, Xiang-Ru Yu and Dae-Ki Kang, "Experimental analysis of na¨ıve Bayes classifier based on an attribute weighting framework with smooth kernel density estimations", DOI 10.1007/s10489-015-0719-1.

15. Yashi Srivastava, Pooja Khanna and Sachin Kumar, "Estimation of Gestational Diabetes Mellitus using Azure AI Services", IEEE 2019.

16. Amita Mahajan, Lois E. Donovan, Rachelle Vallee and Jennifer M. Yamamoto," Evidenced-Based Nutrition for Gestational Diabetes Mellitus", https://doi.org/10.1007/s11892-019-1208-4.

17. Ricardo Augusto Leoni De Sousa, "Gestational diabetes is associated to the development of brain insulin resistance in the offspring", https://doi.org/10.1007/s13410-018-0618-1.

18. Jasmine F. Plows, Clare M. Reynolds, Mark H. Vickers, Philip N. Baker and Joanna L. Stanley, "Nutritional Supplementation for the Prevention and/or Treatment of Gestational Diabetes Mellitus", https://doi.org/10.1007/s11892-019-1199-1.

19. Santosh Kumar and G. Sahoo, "Classification of Heart Disease Using Naive Bayes and Genetic Algorithm", DOI 10.1007/978-81-322-2208-8_25.

20. Md. Maniruzzaman, Md. Jahanur Rahman, BenojirAhammed and Md. Menhazu Abedin, "Classification and prediction of diabetes disease using machine learning paradigm", https://doi.org/10.1007/s13755-019-0095-z.

21. AbidSarwar and Vinod Sharma, "Comparative analysis of machine learning techniques in prognosis of type II diabetes", DOI 10.1007/s00146-013-0456-0.

22. Harry Zhang and Jiang Su, "Naive Bayesian Classifiers for Ranking",Springer 2004.

23. Peter A. Flach and Nicolas Lachiche, "Naive Bayesian Classification of Structured Data", 2004.

24. Yuqian Jiang, Huaizhong Lin, Xuesong Wang and Dongming Lu, "A Technique for Improving the Performance of Naive Bayes Text Classification",Springer 2011.