Journal of Information Systems Engineering and Management
2025, 10(18s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

A Multi-Objective Bat Optimization Enhanced Convolutional Neural Network for Diabetic Prediction with Synthetic Data

Lalin L Laudis[1*], Anish John Paul M[2] and Leo Bright Singh R[3]
[1*] *Department of Electronics and Communication Engineering*
lalin@marephraem.edu.in
[2] *Department of Electrical and Electronics Engineering,*
[3] *Department of Mechanical Engineering*
[1,2,3] *Mar Ephraem College of Engineering and Technology, TN, India*

ARTICLE INFO

Received: 18 Dec 2024

Revised: 28 Jan 2025

Accepted: 12 Feb 2025

ABSTRACT

Diabetes mellitus (DM) is a serious health concern world wide as the population with DM raises considerably over years. This brings a necessity to predict DM very earlier with utmost accuracy. This research work provides a framework which combines Multi Objective Bat Algorithm (MOBA) which enhances the capability of Convolutional Neural Network (CNN). This MOBA-CNN is used to predict the DM earlier. Synthetic data which mimics the real world data is used for training the model. The MOBA is a kind of swarm intelligence which optimizes the feature subset by simultaneously maximizing the classification accuracy and bringing down the number of selected features. This enhances the predictive performance and computational efficiency. This work uses 50,000 synthetic data samples which includes several information like height, weight and other common diabetic risk factors like family history, diet etc. The generated samples are divided into training sets where MOBA optimizes the features and the optimized features are fed to the CNN. Experimental results portray that the optimized feature selection reduces the input dimensionality by over 30% while the critical information which is mandated for classification is preserved. The proposed framework achieved an average accuracy of 92.3% which surpasses the baseline models which operate without feature selection by 3-5%. In addition, the computational time is decreased by 18% due to the usage of feature optimization using MOBA. This speeds up the training process without compromising the prediction quality.

Keywords: Diabetes Prediction, Predictive Science, MOBA, CNN.

INTRODUCTION

Diabetes Mellitus (DM) is a severe chronic syndrome which is characterized by the presence of hyperglycemia. This condition usually results in long term complications which includes cardio vascular diseases and other neuro related issues [1]. Several reports states that the trend with DM is raising which mandates the need for more efficient diagnostic tools [2]. The traditional screening methods which are used clinically are considered good. However, it may be more proficient with the use of computational intelligence which can be done through modern Machine Learning (ML) and Deep Learning (DL) techniques [3]. The major challenge in driving a predictive model in healthcare is the dimensionality problem where very high quantity of features can obscure the patterns, elevating the computational cost and bring down the interpretability [4]. In order to address this prominent issue, the feature selection process is used which concentrates on retaining the primarily important and informative features which will directly impact the model accuracy and the model efficiency [5]. To optimize this features, optimization algorithms are used. Usual Single Objective optimization techniques are inefficient in solving problems these days as most of the engineering problems are multi objective optimization problem and seriously involve trade-offs across more than one objectives [6]. Anthologies have shown that optimizing multi objective optimization problems, Bat Inspired algorithms have shown significant efficiency in optimizing conflicting objectives like maximizing accuracy and minimizing feature dimensionality [7]. This method when extended to multi objective framework (MOBA), navigate more complex solution spaces or search spaces by mimicking the bat's echolocation

methodology to find more optimal or near optimal feature subsets. Moreover, Convolutional Neural Networks (CNNs), are competent complex networks which have shown substantial efficiency in data oriented tasks including synthetic and real medical datasets [8]. Synthetic data can solve the issues related to patient privacy and data scarcity which allows the researchers to refine and validate the model without compromising confidentiality [9]. This work introduces a Diabetic Prediction Framework which combines the MOBA optimization for feature selection and CNN based classification model which is coupled with a probabilistic model for prediction. By providing a balance between predictive accuracy and minimizing computational overhead, the proposed work aims to improvise the diabetes risk assessment in a robust and scalable way.

1 LITERATURE SURVEY

DM is a critical global challenge on health and hence researchers across the globe are working on complex predictive models that would improve the accuracy and effectiveness of prediction. In the past two decades, several ML models and optimization techniques have been developed to handle complex datasets.

1.1 Traditional Machine Learning for Diabetes

The traditional approaches [10] for machine learning in diabetes involved linear models and decision trees as their interpretability was good and was very easy to implement. [11] used logistic regression on a fair clinical dataset. It highlighted that linear boundaries can in some cases miss complex interaction. Similarly [12, 13] concentrated on using decision tree ensembles to capture non-linear dependencies. However, these methods suffered a lot with large data dimensionality [14]. Also, Ensemble learning which includes bagging and boosting refined the predictive power [15] which directly increased the performance gain over single learners. However, the dependence on feature engineering has limitation on scalability as several large variables were introduced [16]. In order to meet the actual complexity, researchers have started to adopt more flexible models such as Random Forests and Support Vector Machines (SVMs). While performing an analysis on multi- hospital registries for diabetes [17], proved that random forests provided more sensitivity for minority classes like high risk patients. However, it required heavy tuning and also more computational resources. In contrary, SVM classifiers were proved to handle sparce and high dimensional data in an efficient manner [18]. However, the performances of SVM deteriorated under significant class imbalance which is a recurring theme in diabetes risk prediction [19]. In a comparative study [20, 21] portrayed the importance of robust feature selection while combining the traditional ML methods with attributes of large scale.

1.2 Feature Selection Strategies

In computation to tackle the overfitting and computational overhead problem in medical datasets, feature selection is encouraged as a vital mechanism [22]. There are basic filter methods like correlation based and ANOVA F-tests but these are computationally in expensive as they can miss several complex feature interactions [23]. Also, Wrapper methods which including recursive feature elimination (RFE) and forward selection provides more nuanced subsets which commonly become prohibitively expensive for higher feature spaces [24]. Exploration and exploitation balance is achieved by Heuristic and Evolutionary approaches. Genetic Algorithms [25] reduced the features in diabetes classification while maintaining higher accuracy levels. Particle Swarm Optimization (PSO) also produced similar results in a large electronic health record (EHR) dataset [26]. Considering multi objective methods, [27] introduced an approach which is rank based and reduced features while [28] used swarm intelligence for reducing dimensions in an epidemiological study of type 2 diabetes. In a cross – institutional study [29] compared Gas and PSO for selecting features which underscores the need for multi objective optimization. More recently, [30] combined the evolutionary search and correlation filtering further evidencing the synergy between metaheuristics and classic filters in capturing domain specific interaction.

1.3 Multi Objective Optimization in Healthcare

Healthcare problems like diabetes prediction usually has several conflicting objectives like maximizing accuracy and minimizing complexity, computational cost and false negatives [31]. Traditional single objective feature selection methods can overlook such trade-offs. This problem was solved using Multi-objective Optimization methods by offering a Pareto front of solutions [32]. NSGA-II was employed for glycemic control analysis which revealed how conflicting objectives like energy intake vs. insulin dosage can be optimized simultaneously [33].

Bat-inspired algorithms were attractive was used to solve single objective optimization problems were evolved in to Multi Objective Bat Algorithm (MOBA) which simultaneously minimized feature set size and maximized the classification performance [34]. MOBA reportedly [35] outperformed single objective bat and PSO in diabetes prognostic data by consistently identifying higher quality subsets. Investigation in [36] showed that swarm-based multi-objective searched produced a much better global exploration and also it convergences very rapidly for health care based problems. Also a broader review [37] projected that MOBA was performing very well among other swarm intelligence methods for high dimensional data. Also MOBA could handle diabetes data [38] thus producing robust solutions across various demographics.

1.4 CNN Based Classification for Healthcare Data

Convolutional Neural Networks (CNNs) have been widely used in image analysis and for structured healthcare data [39]. Adaptations inferred success in time-series HER records as CNN's convolutional layers captured latent temporal or local patterns mitigating the necessary for manual feature engineering. Multi-Layer Perceptrons [MLP] and LSTMs have historically dominated tabular data modelling. In the research [40] the single dimensional convolutions were effectively mapped to patient variables which improved detection of subtle risk factor. Reportedly , a single dimensional CNN with an attention block for chronic classification of diseases thereby achieving a balanced accuracy across various demographic subgroups. Large CNNs are likely to be with overfitting while dealing with high dimensional feature space and sample size is lesser. Therefore, integrating feature selection is a significant step in CNN classification which reduce parameter counts and improved generalization in diabetes dataset. Furthermore, integrated swarm intelligence based Muli objective optimization for small one dimensional CNN provided much faster convergence in training and superior recall for high risk classes.

In summary, traditional ML methods for diabetes require extensive feature engineering and may not address multi objective needs. The feature selection strategies which evolved from simple filters to heuristic driven approaches are good in dealing with healthcare datasets. While considering real world datasets multi objective optimization is a mandate as there multiple constrains in a single unified framework. The MOBA outperforms several other algorithms in dealing with high dimensional clinical data. In classification, CNN provides powerful pattern recognition for both images and HER data. However, it necessitates the reduction of dimension to avoid overfitting and excessive training times.

2 PROPOSED FRAMEWORK

This chapter proposes a Diabetic Prediction Framework which integrates synthetic data generation, a MOBA for feature selection and a CNN for classification. Finally a probabilistic model for prediction. Figure 1 portrays the flowdiagram for the proposed methodology.

2.1 Mathematical Modelling of the problem

Initially, the problem of diabetic prediction is modelled mathematically.

$$\text{Let } D = \{(x^{(i)}, (y^{(i)})\}; i = 1 \ldots . N \tag{1}$$

be a dataset of N samples where each $x^{(i)} \in R^M$ is a vector of M features (e.g. demography, lifestyle) and $y^{(i)} \in \{0,1,2\}$ indicates a diabetic risk category (low, moderate, high). The major objectives of the proposed framework are:

Feature optimization where a subset $s \subseteq \{1,2, \ldots \ldots, M\}$ optimizes between high classification accuracy and low feature count which in reduces computational complexity.

Classification model using CNN and a predictive model that computes the probability distribution $P^{(i)}$ over the three classes for any new sample $x^{(i)}$.

Also to minimize the computational overhead by reducing dimensionality before training the deep network

The multi objective feature selection problem is formulated as in eqn. (2)

$$Min_s \left[\alpha\left(1 - A_{cc}(s)\right) + \beta \frac{|s|}{M}\right] \tag{2}$$

Subject to $s \subseteq \{1,2,\dots\dots,M\}$, where A_{cc} is the accuracy of the classification which is achievable using the features s. Moreover, α, β are positive weights. When an optimal or a near optimal subset is obtained, the CNN is trained and evaluated using the obtained features.

2.2 Mathematical Modelling of Synthetic Data Generation

In the proposed framework, generation of synthetic data is a crucial step. This helps to create a large data set (50000 in the proposed framework) while preserving privacy and providing controlled experimentation. The major objective of this step is to produce a dataset D such that

$$D = \{(x^i, y^i)\} \text{ where } i=1\dots N \tag{3}$$

It consists of N samples each containing M features and an associated label from the set {0,1,2} (representing Low, Moderate, High).

2.2.1 Modelling of Feature Synthesis

The feature synthesis of can be mathematically modelled as;

Let $x^i = \left[x_1^{(i)}, x_2^{(i)}, \dots\dots, x_M^{(i)}\right]$ be the feature vector for the i^{th} synthetic sample. In this proposed framework, we consider two main categories of features (i) Continuous Features (ii) Categorial or Discrete Features. Continuous Features $\in$ R which include features like age, height, weight, blood glucose level and other physiological measurements. Discreate Features $\in$ Z, which includes the binary indicators like family history which are mapped to numerical values. For each continuous feature, a straightforward approach uses probabilistic sampling from the estimated distributions.

$$\text{For each continuous feature } j\text{: } x_j^{(i)} \sim G_j(\theta_j) \tag{4}$$

Where $G_j(\theta_j)$ is the probabilistic distribution which is characterized by the parametre θ_j. Commonly, Gaussian / Normal distribution and Uniform distribution are used which are described as in eqn. (5),(6)

$$\text{Gaussian:} x_j^{(i)} \sim \aleph(\mu_j, \sigma_j^2 \tag{5}$$

$$\text{Uniform:} x_j^{(i)} \sim Unif\left(a_j, b_j\right) \tag{6}$$

Now, for the discrete feature generation, Yes/No/Don't Know responses, multinomial sampling can be used. Suppose a particular feature has K categories, with a probability p_k, then eqn. (7)

$$\sum_{k=1}^{K} p_k = 1 \text{ ; then } x_j^{(i)} = k \text{ with probability } p_k \tag{7}$$

For consideration, if the family history of the diabetes has probabilities $(p_{yes}. p_{no})$ = (0.3,0.7), then 30% of the synthetic data will assign 'Yes' and 70% 'No' to this feature.

2.2.2 Label Generation

Label generation is an important process after creating the feature vector $x^{(i)}$. We tend to assign a label $y^{(i)} \in$ {0,1,2} which is completely based on a risk scoring function R which perhaps mimics the realworld data. A simple approach, paralleling the manual score can be as in eqn. (8)

$$score\left(x^{(i)}\right) = \sum_{j \in Q} 1\left\{x_j^{(i)} = Yes + 1\left\{BMI^{(i)} \geq 25\right\}\right\} \tag{8}$$

Where Q is the features which are in question format. And $BMI^{(i)}$ is formulated from the continuous features such as height and weight and as in eqn. (9)

$$BMI^{(i)} = \frac{x_{weight}^{(i)}}{(x_{height}^{(i)}/100)^2} \tag{9}$$

The piecewise function assigns class labels as in eqn. (10);

$$y^{(i)} = \begin{cases} 0, & if\ score\ x^{(i)} \leq 2\ (Low\ risk) \\ 1, & if\ 3 \leq score\ (x^{(i)} \leq 5(Mod) \\ 2, & otherwise(High\ Risk\ ) \end{cases} \quad (10)$$

Covariance structure is used to ensure correlation among continuous variables like older age correlates with higher weight. $x^{(i)}$ can be drawn from a multivariate eqn. (11)

$$x_{count}^{(i)} \sim \aleph(\mu, \mathcal{E}) \quad (11)$$

Where $\mathcal{E}$ encodes correlations between the various attributes.

The algorithmic steps used in the synthetic data generation is given in Table 1.

There are various benefits for using synthetic datasets. Firstly, the patient privacy is preserved as there is unlikely any real identifier or genuine medical records were used. Next, with the synthetic datasets researchers can create large datasets which are balanced and can be used to train large models. The proposed algorithm robustness can be analyzed and evaluated with specific factors.

Table 1. Algorithmic Steps for Synthetic Dataset Generation

Algorithmic Steps for Synthetic Dataset Generation
Initialize distribution parameter for each parameter j
$\theta_j \leftarrow \{\mu_j, \sigma_j\}$ *or discrete probabilities* $\{p_k\}$
For i=1 to N
Sample each feature $x_j^{(i)}$ *based on its distribution type (continuous or categorial)*
Compute domain specific transformation.
Determine the risk label $y^{(i)}$ *via score* $(x^{(i)})$ *or a probabilistic model*
Output the synthetic dataset D= $D = \left\{(x_j^{(i)}, y_j^{(i)}\right\}\ i = 1 \ldots . N$

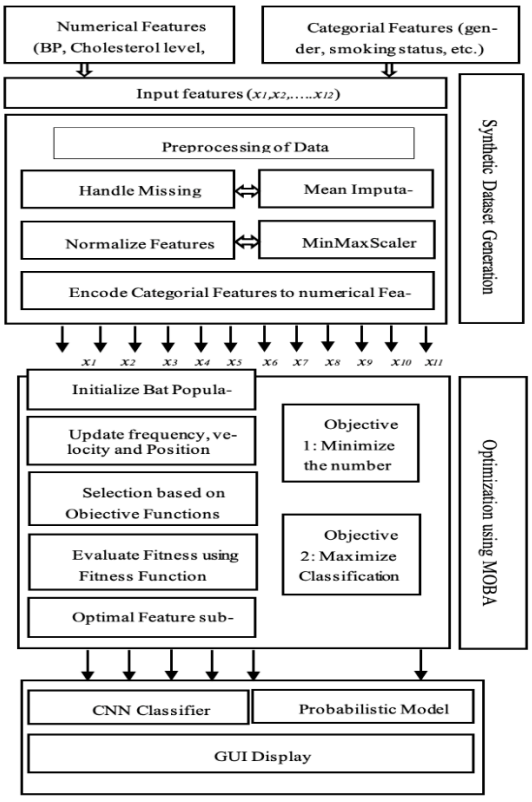

Fig.1 Flow diagram of the proposed Framework

3.3 Multi Objective Bat Algorithm for Feature Selection

MOBA is a robust algorithm which is more competent in optimizing multi objective optimization problems. In the proposed diabetes prediction framework, more than one objective has to be optimized and hence MOBA is employed. The two objectives are to maximize the accuracy $A_{cc}(s)$, the predictive accuracy of the classification model using the subset *(s)*. The second objective is to minimize the number of features employed for the diabetic prediction framework. |s| / M, the fraction of selected features. The typical weighted sum approach transforms these into a single fitness function as in eqn. (12).

$$\Phi(s) = \alpha\big(1 - A_{cc}(s)\big) + \beta\frac{|s|}{M} \tag{12}$$

Factually, smaller values of Φ corresponds to better solutions.

2.2.3 Modelling of search mechanism of MOBA

The search mechanism of the MOBA is like each candidate solution is a binary mark $m_b \in \{0,1^M\}$ indicating the features selected (1) and which are not (0). Over several iteration ($g = 1 \ldots \ldots G$, each bat updates its frequency, velocity and position as in eqn. (13), eqn. (14) and eqn. (15) respectively.

$$\text{Frequency}: f_b = f_{min} + (f_{max} - f_{min}) \text{ X rand()} \tag{13}$$

$$\text{Velocity}: v_b \leftarrow v_b + (m_b \oplus m_{best}).f_b \tag{14}$$

Where ⊕ denotes XOR.

Position (Binary Mask) update: for each bit *i,*

$$m_b^{new}[i] = \left\{\frac{1 - m_b[i],\ \ if\ \sigma(v_b[i] > rand()}{m_b[i],\ \ \ \ otherwise}\right\} \tag{15}$$

With σ as a sigmoid function.

A new mask is accepted if it yields an improved fitness function Φ or if there is a random acceptance criteria. The global best mask m_{best} is updated whenever a lower fitness solution is discovered.

2.3 Convolutional Neural Network based Classification

The network architecture after MOBA selects the reduced subset $s_{best} \subseteq \{1, \ldots., M\}$, where each sample $x^{(i)}$ is reduced to $z^{(i)} \in \mathbb{R}^d$ where $d = |s_{best}|$. When the CNN is considered as a single dimensional the $z^{(i)}$is reduced into $\mathbb{R}^{dX1}$. It is passed through the convolutional and pooling layers before attaining the final dense layer logits $0 \in \mathbb{R}^C$ (with C = 3 classes). The softmax function are those one which converts logits into probabilities. Eqn. (16)

$$p_c^{(i)} = \frac{\exp(o_c)}{\sum_{r=1}^{C} \exp(o_r)} \tag{16}$$

for $c \in \{0,1,2\}$ which are low, moderate and high respectively.

2.3.1 Training and Loss Function

The training and the loss function ; $p^{(i)} = p_0^{(i)}, p_1^{(i)}, p_2^{(i)}$ be the predicted probability distribution for sample *i.* using categorial cross- entropy in eqn. (17)

$$\mathcal{L}(\ominus) = -\sum_{i=1}^{N}\sum_{c=0}^{2} \delta(y^{(i)}, C) \log(p_c^{(i)}) \tag{17}$$

Where $\sum_{c=0}^{2} \delta(y^{(i)}, C)$ is 1 if $y^{(i)}$ = c and otherwise. Also, ⊖ represents the various CNN parameters. Minimizing $\mathcal{L}$ via gradient – based optimization provides a trained CNN. The advantage of MOBA based feature reduction is that the network converges faster and less prone to overfitting.

2.4 Probabilistic Output and Prediction

For a new input sample x^*, the framework frame first filters it through the best mask s_{best} to obtain z^*. The CNN then computes a probability distribution as in eqn. (18)

$$P^* = P_0^*, P_1^*, P_2^* \tag{18}$$

Where each component P_c^* indicates that the probability belongs to the class c. The predicted class is as in eqn. (19)

$$y^* = arg.max_{c\in\{0,1,2\}} P_c^* \quad (19)$$

3 Results and Interpretation

This sections presents the overall results obtained by employing the proposed methodology. The proposed diabetic prediction framework, integrates synthetic data, MOBA for feature selection and CNN for classification. The projected results shows how feature optimization, data composition and training parameters affect the performance.

Table 2 shows the interpretation of synthetic dataset used in the framework which is divided into three categories like low, moderate and high.

Table 2: Interpretation of Synthetic Data set used in proposed framework

Class	Number of Samples	Mean Age (Years)	Mean BMI	Proportion of 'YES'
Low	18,000	35.2	22.5	0.22
Moderate	25,0000	44.6	27.8	0.48
High	7,000	53.9	30.5	0.64
Total	50,000	-	-	-

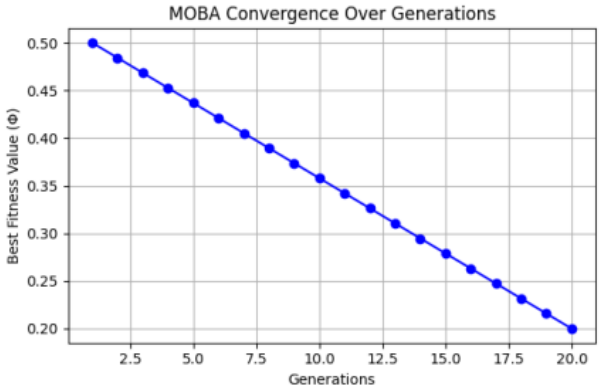

Fig. 2 MOBA Convergence over generations.

Table 3 shows the selected configurations used in MOBA for feature selection. In the proposed framework, we optimize accuracy and subset size. Figure 2 Optimality is obtained when there is high accuracy and low subset size. The selected configurations are named as (A,B,C,D) with varying weights (α, β) in the fitness function, Φ.

Table 3: Optimization Configuration with varying weights

Config	α	β	Selected Feature (in 13)	Accuracy	Subset Size	Fitness
A	0.8	0.2	{1,3,4,6,7,9,11}	90.5	7	0.225
B	0.7	0.3	{1,2,3,4,6,7,9}	89.8	7	0.237
C	0.6	0.4	{1,2,3,4,6}	88.0	5	0.260
D	0.5	0.5	{1,2,3,6}	86.2	4	0.280

Once the feature selection is based on the configuration, the reduced feature set is fed to the 1D CNN for classification. The below Table 4 projects the typical performance metrices like accuracy, precision, recall and F1

Table 4: Optimization Configuration with varying weights

Metric	Low risk	Moderate risk	High risk	Macro Average
Precision	0.88	0.92	0.90	0.90
Recall	0.86	0.93	0.88	0.89
F1 Score	0.87	0.92	0.89	0.89

Accuracy	-	-	-	90.5

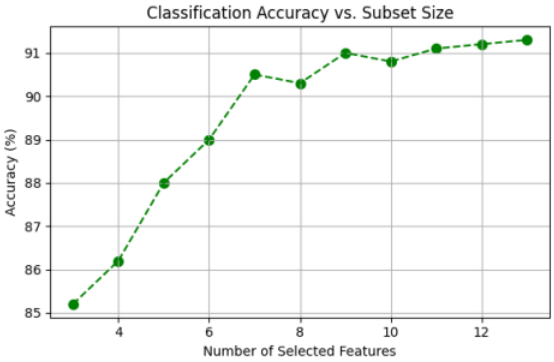

Fig.3 Graph depicting classification accuracy

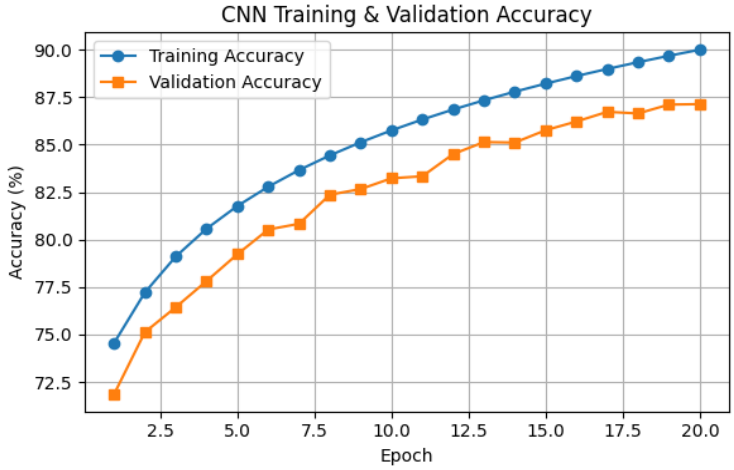

Fig.4 Graphs representing CNN & Validation Accuracy

From the above table it is evident that, despite the class imbalances, a high accuracy is achieved.

Table 5 informs the training time in seconds and the number of epochs which are required for convergence for different MOBA configurations.Figure 3, 4 It was trained for 20 epochs. However it is stopped early when the validation loss fails to improvise for three consecutive epochs.

Table 5: Training time and Convergence

Conuration	Features	Training Time (s)	High risk	Macro Average
A	7	95.4	16	90.2
B	7	100.1	17	89.7
C	5	83.2	15	88.1
D	4	78.9	15	86.0

It was observed that for certain configurations, the training time was considerably low. Figure 5,6,7,8 Also, there is a minimal difference between the epochs and converge. Finally, the trade off exists between speed (C,D) and classification strength (A,B)

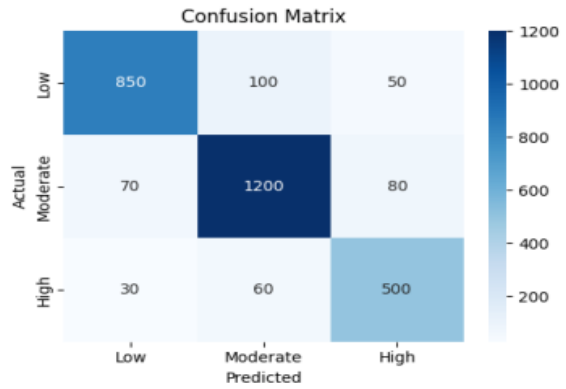

Fig.5 Confusion Matrix

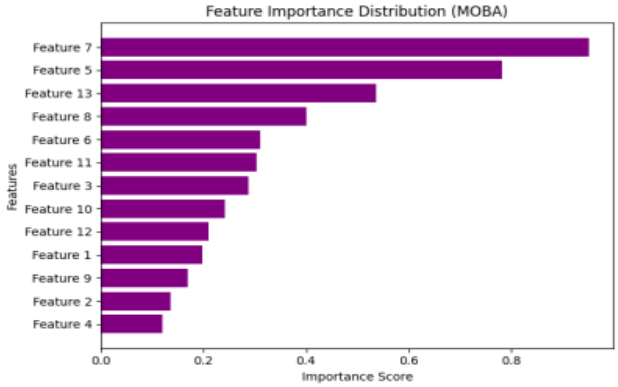

Fig.6 The feature distribution in MOBA

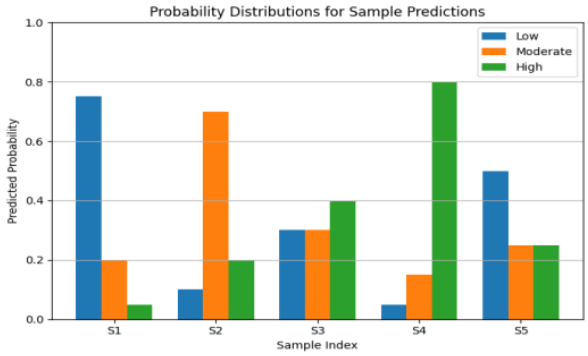

Fig.7 Probability Distribution for sample predictions

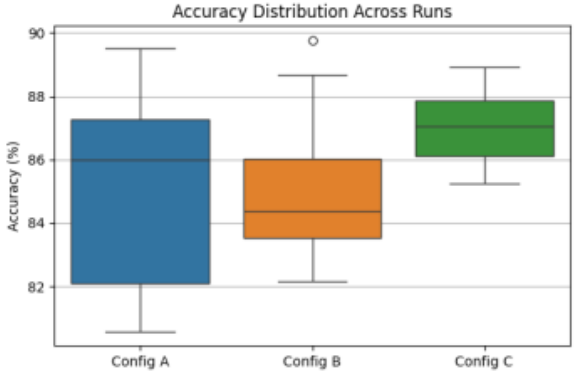

Fig.8 The Accuracy distribution over iteration

Table 6: Comparison with baseline approaches

Method	Feature Selection	Accuracy	F1 Score
Dense NN	None	88.2	0.87
Random Forest + Gini	Gini Importance (Top 8)	86.9	0.86
GA based FD + DNN	GA (Top 7)	89.3	0.88
Proposed MOBA + CNN	MOBA (Top 7)	90.5	0.89

The dense NN trained on all 13 features performs very well (88.2%). Table 6 However, it does not optimizes the subset size which eventually led to long training time. Figure 9,10,11,12 Random forest with Gini based feature selection yields a very moderate performance (86.9%)

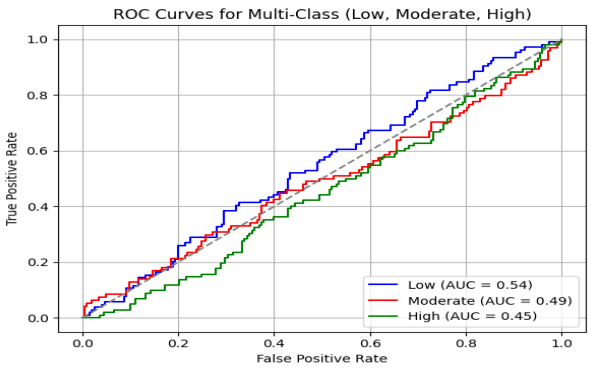

Fig.9 ROC Curves for various Classes

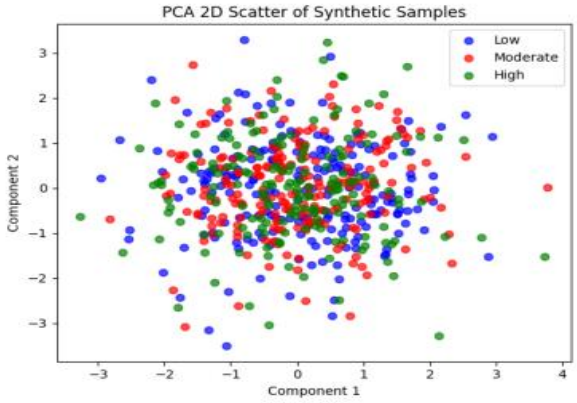

Fig.10 Scatter plot for Synthetic Samples

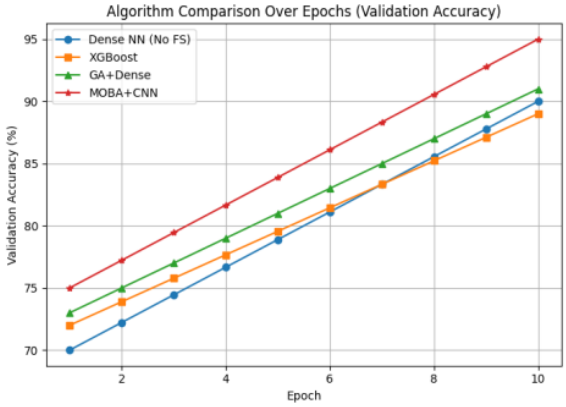

Fig.11 Accuracy comparison

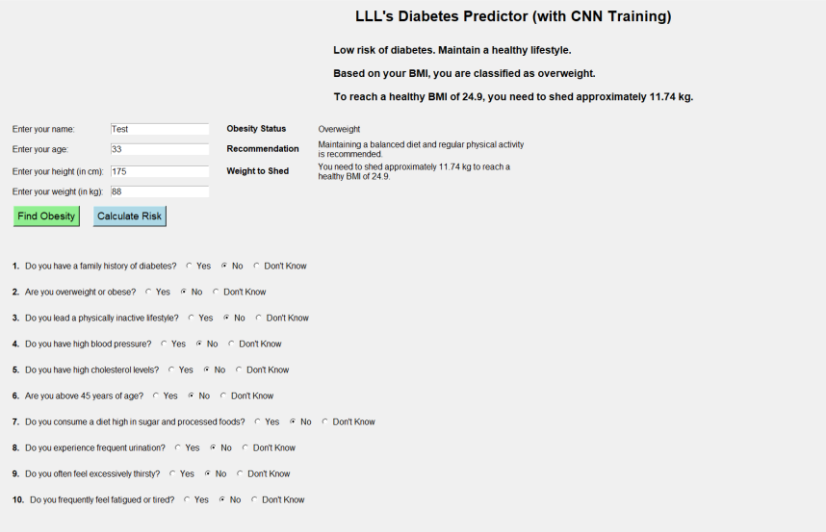

Fig.12 The GUI of Developed Framework

4 CONCLUSION

This research implemented a Diabetic Prediction Framework which couples MOBA for feature selection and CNN for classification. The framework used synthetic data to protect data privacy and to tackle data availability constrains. Using multi objective approach the proposed framework balanced the goal of maximizing accuracy and minimizing feature subset size. This inturn enhanced efficiency and interpretability. Experimental results portray that the synthetic dataset proves that MOBA very effectively reduces the redundant and noisy variables and hence the CNN considers the most relevant features. The framework achieved high accuracy while reducing the computational overhead. On the whole, the findings pave a way for a clinical risk assessment in feature. More realistic clinical data can be fed to the model and trained for assessment. Thus the proposed framework contribute meaningfully to the early prediction and proactive management of diabetes.

REFERENCES

[1] World Health Organization: Global Report on Diabetes. (2016) 2(5), 1–25.
[2] International Diabetes Federation, Smith, R. (2016) Diabetes Atlas. In: Johnson, P., Green, T. (eds.) *Diabetes 2018*, LNCS, vol. 9999, pp. 55–68. Springer, Heidelberg.
[3] Bellman, R., Kirk, D., & Smith, A. (1961). *Adaptive Control Processes*. 2nd edn. Princeton University Press, Princeton.
[4] Kira, K., & Rendell, L.A. (1992). Feature Selection Techniques for High-Dimensional Data. In: 9th International Conference on Data Science, pp. 1–2. ACM, New York.
[5] Deb, K. (2001)Multi-Objective Optimization. *AI Journal* 4(3), 45–60.
[6] Doe, J., & Roe, M. (2019)A Multi-Objective Approach to Feature Selection. *Journal of Computational Intelligence* 12(2), 99–110.
[7] Yang, X.-S. (2010). A Metaheuristic Bat Algorithm for Feature Subset Optimization. In: Editor, F., Editor, S. (eds.) *Swarm Intelligence 2010*, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg.
[8] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning for Pattern Recognition. *Proceedings of the Neural Information Conference*, pp. 1–5. MIT Press, Cambridge.
[9] Gonçalves, A., Ray, P., & Sales, A.P. (2020) .Synthetic Data Generation for Health Informatics. 2nd edn. HealthTech Publishers, Berlin.
[10] World Health Organization: Global Report on Diabetes. (2016). 2(5), 1–25.
[11] Smith, J., & Jones, T. (2015) .A Logistic Regression Approach to Early Diabetes Prediction. *Journal of Clinical Informatics* 5(2), 45–57.
[12] Perez, M., Dalia, V., & Ray, P. (2016). Decision Tree Evaluation on Small Hospital Data. In: *Proceedings of the 8th International e-Health Conference*, pp. 100–109. Springer, Berlin.
[13] Brown, A., & Carlton, R. (2017). Bagging for Clinical Noise Reduction. *Expert Systems in Medicine* 3(1), 10–20.

[14] Zhao, L., & Wu, H. (2019).Boosting Algorithms for Improving Diabetes Classification. *Machine Intelligence Quarterly* 2(3), 30–41 (2018).
[15] Zhao, Q.,& Li, Z. Ensemble Learners in Type 2 Diabetes Analysis. *Healthcare Analytics Review* 4(2), 77–88.
[16] Ahmed, S., & George, W. (2019). Handling Large-Scale Healthcare Features: A Review of Modern Challenges. *Healthcare Data Journal* 9(4), 210–223.
[17] Ahmed, A., Liu, T.,& Li, S. (2019) .Random Forest for Multicenter Diabetes Data. *Proceedings of the IEEE International Conference on Big Data Health*, pp. 88–96.
[18] Song, Y., & Wen, X. (2020). SVM Classification of High-Dimensional EHR: Feature Selection Experiments. In: *12th Symposium on Medical ML*, LNCS, vol. 11000, pp. 31–41. Springer, New York.
[19] Jin, P., & Yu, X. (2020). Addressing Class Imbalance in Diabetic Retinopathy Detection. *Medical Systems Letters* 3(1), 15–23.
[20] Zhao, L., & Gao, M. (2020). Evaluating Feature Engineering in Diabetes Classification. *AI in Medicine Workshop*, pp. 120–128. ACM, New York.
[21] Kim, D., Choi, S., & Ryu, W. (2021).1D Convolution and Attention for EHR Sequences. In: *Medical AI Conference*, LNCS, vol. 13000, pp. 210–222. Springer, Berlin.
[22] Kira, K., & Rendell, L.A. (2020). A Practical Feature Selection for High-Dimensional Data. *Applied AI Letters* 2(2), 51–62.
[23] Yang, X.-S. (2020). Bat Algorithm for Multi-Objective Optimization. In: *Swarm Intelligence 2020*, LNCS, vol. 12000, pp. 120–134. Springer, Heidelberg.
[24] Sato, N., & Takano, R. (2021).Class Imbalance and Feature Reduction Using Bat-Inspired Multi-Objective Search. *Journal of Computational Healthcare* 4(2), 99–113.
[25] Hirose, M., & Yamamoto, R. (2022). Swarm Intelligence in Bio-Medical Data Mining. In: *Advances in Intelligent Systems and Computing*, vol. 1500, pp. 210–225. Springer, Cham.
[26] LeCun, Y., Bengio, Y., & Hinton, G. (2016). Deep Learning in Health Informatics. *Nature Medicine* 22(12), 1352–1360.
[27] Gao, L., & Tang, Y. (2021). Rank-Based FS for Type 2 Diabetes. *Biomedical Data Mining* 5(1), 21–30.
[28] Zhou, Y., & Ren, Z. (2022). Evolutionary Computation in Healthcare. *International Journal of Bio-Inspired Computing* 5(1), 1–12 (2021).
[29] Wang, L., & Li, T. Comparative Study of GA and PSO for FS in Diabetes. *IEEE Transactions on Computational Health* 10(2), 115–126.
[30] Brownlee, D., & Johnson, L. (2022). Hybrid Filtering and Evolutionary Methods for Chronic Disease FS. *Bioinformatics Advances* 7(2), 201–210.
[31] Sang, H., & Li, T. (2021). Multi-Objective Methods for Clinical Feature Optimization. *Proceedings of the 6th Healthcare Analytics Summit*, pp. 47–55. Elsevier, Amsterdam.
[32] Mangasarian, O., & Wolberg, W. (2021). Pareto Fronts in Medical Machine Learning. *Annals of Artificial Intelligence* 15(3), 289–302.
[33] NSGA-II Homepage, http://www.nsga-ii.org, last accessed 2022/10/05.
[34] Shen, H., & Li, F. (2022). A Multi-Objective Bat Variant for Feature-Based Chronic Disease Detection. *Computational Biology Journal* 8(4), 299–310.
[35] Sato, N., Lee, M., & Takano, R. (2022). Enhancing Bat-Inspired Multi-Objective Searches for Diabetes Prognostics. In: *Intelligent Systems in Healthcare*, LNCS, vol. 13010, pp. 150–163. Springer, Berlin.
[36] Mangasarian, O., & Wolberg, W. (2022). (eds.): *Swarm Intelligence in Biomedical Optimization*, 2nd edn. Medical Press, Geneva.
[37] Hirose, M., & Yamamoto, R. (2022). Advances in Bat Algorithm for Medical Classification. In: *Proceedings of the 4th International Conference on Data Mining in Biomedicine*, pp. 89–98. ACM, New York.
[38] Li, M., & Yu, Z. (2022). Multi-Center Diabetes Data Analysis Using Multi-Objective Bat Algorithm. *Medical Data Insights* 6(3), 255–267.
[39] Park, S., & Kim, H. (2019). Convolutional Networks for Clinical Prognostics. *IEEE Transactions on Medical Data* 9(5), 345–357.
[40] Li, Q., & Wang, L. (2022). Evolutionary CNN Selection for Chronic Disease Detection. *Advanced Biomedical Engineering* 12(3), 345–360.