**Paper:**

# Ensemble Learning-enabled Type 1/Type 2 Diabetes Classification through Flash Glucose Monitoring Data

**Yicun Liu, Wei Liu, Xiaoling Cai, Rui Zhang, Zhe An, Dawei Shi, Linong Ji**

State Key Laboratory of Intelligent Control and Decision of Complex Systems, School of Automation, Beijing Institute of Technology, Beijing, China
E-mail: liuyicun@bit.edu.cn, daweishi@bit.edu.cn
Department of Endocrine and Metabolism, Peking University People's Hospital, Beijing, China
E-mail: liuwei850217@163.com, dr june@sina.com, rachelhope@126.com, jiln@bjmu.edu.cn
Global Energy Interconnection Research Institute Co., Ltd.
E-mail: jilla1@163.com

**Abstract. Clinically, diabetes type diagnosis is difficult and inaccurate by relying on biochemical indicators and doctors' experience. In order to improve the efficiency and accuracy of diabetic diagnosis, we propose a data-driven diabetes type classification algorithm, called ensemble-based linear discriminant analysis (LDA) algorithm. Furthermore, supported by the Department of Endocrinology And Peking University People's Hospital, a dataset containing blood glucose records of 113 diabetic patients is created for model training and testing. In our algorithm, to reduce the data redundancy and increase the category discrimination, a data preprocessing method composed of data reorganization and downsampling is presented, in which the classification accuracy is significantly improved by 22% on average. The LDA method is adopted to build a generalized base model. Then an ensemble model is formed by integrating the LDA and the data preprocessing methods utilizing stepwise functions. In the experiment, compared with five machine learning algorithms and their combination with the Adaboost algorithm, our proposed algorithm obtains the highest sensitive value (0.8182) in diabetes type classification and the largest F-Measure, Matthews correlation coefficient score (81.82%, 72.73%). The model achieves satisfactory classification performance in terms of the percentage of classification accuracy which is 87.88%.**

**Keywords:** Ensemble learning, Data processing, Linear discriminant analysis, Diabetes types classification, Machine learning.

## 1. Introduction

Diabetes is the third leading cause of death following diseases of heart and cancer. The estimated number of people with diabetes for the age group 20-79 years was 415 million (uncertainty interval: 340-536 million) in 2015 and is expected to reach 642 million (uncertainty interval: 521-829 million) by 2040 [22]. The total number of deaths attributed to diabetes is estimated to be 5 million in people with diabetes for the age group 20-79 years [22]. Compared with the normal population, diabetes patients have higher risk of being infected, which potentially increases their morbidity and mortality [5]. Therefore, accurate diagnosis of diabetics as early as possible is essential to accelerate diabetics recovery and reduce physical damage.

According to the American Diabetes Association (ADA) diagnoses criteria [1], diabetes could be diagnosed based on A1C criteria or plasma glucose criteria. But diabetes diagnosis is quite a challenging task for medical practitioners due to complex interdependence on various factors. Thus, more and more data-driven diseases classification and prediction approaches are presented in [2, 7, 11, 14–16, 25–28]. The electronic health record was used for diabetes prediction and diagnoses [24]. But, these methods required long-term multiple types of medical data recording. In order to improve the classification efficiency of diabetes, direct classification of type 2 diabetes from retinal fundus images methods were proposed in [13]. Furthermore, Jake *et al.* utilized combining elemental analysis of toenails and machine learning techniques as a non-invasive diagnostic tool for the robust classification of type-2 diabetes [4]. Rich biological information data helped predict and diagnose diabetes [10, 12]. However, the difficulty and complexity of data acquisition also increased. To achieve a conducive rapid diabetes diagnosis, Ashenafi *et al.* adopted the blood glucose data for type 1 diabetes classification and tested the state of art machine learning methods [3]. Kannadasan *et al.* also employed a combination of conditional decision-making and supported vector machine approach based on blood glucose to analyze and predict type 2 diabetes [17]. More machine-learning-based approaches for diabetes prediction and classification were proposed in [8, 9, 18, 32].

In Asia, the development of workplace careers is closely related to their health. Detailed classification of diabetes is one of the important health evaluation indicators. But the algorithms mentioned above concentrate on a single type of diabetes classification and prediction research, and cannot distinguish between type 1 and type 2

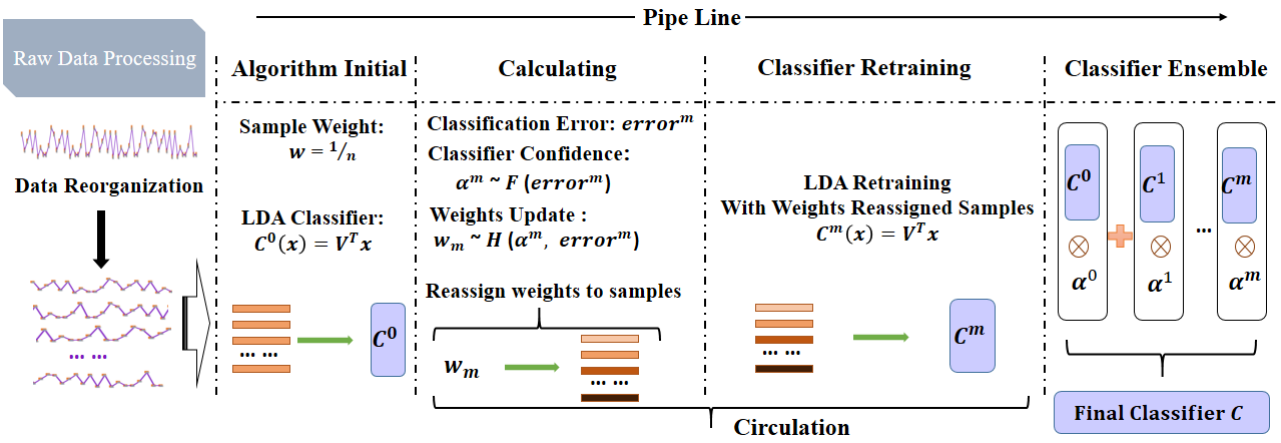


**Fig. 1.** The diabetes type classification model framework diagram.In the figure, the dense sequence data at the far left column indicates the original blood glucose sample data of unequal length, which is generated as a new data sample of equal length after data preprocessing. The new sample is represented by a solid rectangle in the figure. The darker the color, the greater the weight of the sample in the resampling process. The purple square represents the classifier, and the middle weak classifier generated by the model is indicated by the letter $C$ and the superscript $m$.

diabetes at the same time. However, there is no significant difference between type 1 and type 2 diabetes in blood glucose data. In our study, it is difficult to construct a huge dataset with blood glucose data, and the data features dimension is sparse. Due to the limited fitting and feature representation capabilities of a single model on a sparse dataset, those facts direct affect the single model classification performance. More importantly, the generalization ability of a model is an important indicator to measure the effectiveness of a model. Single models are usually susceptible to incomplete data, which leads to overfitting and underfitting. Such that, a single model is difficult to achieve generalization and solve practical problems.

Considering the single model limitations, ensemble learning has better generalization ability and features expression. It promotes the single model classification performance by the idea of weighted integration. For instance, John *et al.* utilized multi-source ensemble learning to promote the feature representation capabilities with incomplete data. [23]. Tao *et al.* develop an evolutionary ensemble modeling to improve the accuracy of warfarin dose prediction [29, 30]. For further results based on ensemble learning classification strategy, see [19–21] and references therein.

Therefore, in this study, a new algorithm motivated from ensemble learning for type 1 and type 2 diabetes classification is proposed through a blood glucose data record. The main contribution of this work is listed as follow:

(a) To reduce the redundancy of irrelevant data and improve distinguishable of category features, a data preprocessing method is proposed including data reorganization and down-sampling. Furthermore, consideration of a data visualization experiment for the processed dataset towards feature analysis;

(b) Fusion ensemble learning strategy and Linear Discriminant Analysis algorithm are proposed , in order to distinguish detailed diabetes type. Compared with the present machine learning algorithms, including Naive Bayes (NB), Logistic Regression (LR), K-Nearest Neighbors (KNN), Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM), our method achieve competitive results for diabetes type classification.

## 2. Methods

In this section, the diabetes type classification algorithm is proposed. As illustrated in Fig. 1, the final diabetes type classifier is obtained through raw data preprocessing, weak classifier pre-training, cyclic retraining and sample weighted resampling based on the classification error.

### 2.1. Flash Glucose Monitoring Data Preprocessing

The original data set is composed of FGM data continuously collected over a while at a fixed frequency. Because each participant has a different length of time participating in the project, the length of data recorded by each independent sample is different. Moreover, the change in human FGM data is a sluggish and long-term process. Therefore, the FGM data is susceptible to the sudden behavior of individuals, such as irregular eating or irregular sports behavior at a certain stage or random errors of sensors. Those happenings would reduce the data quality, and increase the data features extraction difficulty. To solve the inconsistent sample data length, reduce the impact of individual differences and the sensor random errors, we propose a data preprocessing method including data reorganization and down-sampling. First, we segment each independent data sample in units of days to produce a new subset of samples that obey the same distribution and share the same data label. Each sample in the subset is

equal in length, and the starting point is to sample at 0 o'clock every day. Second, down-sampling the new data samples in each subset. Finally, the down-sampled data is stitched along the time series to synthesize into a new sample set and keep the data length of each sample in the new sample set consistent before down-sampling. In addition, down-sampling FGM sequences also bring additional accuracy improvements, which will be discussed in the section IV.

### 2.2. Linear Discriminant Analysis

LDA is a type of feature compression learning algorithm. The basic idea of LDA is to project the features in higher dimension space into a lower dimension space to achieve the effect of extracting classification information and compressing the feature space dimension. In this paper, LDA is applied as a weak classifier, which separates the FGM data into two classes to distinguish the diabetes type. LDA explicitly attempts to model the difference between the classes of data by calculating the within-class and between-class scatter matrices. Then the algorithm starts with a projecting, which yield optimum separation between the groups. The input data is projected into LDA space according to equation (1), and the projecting form is described as follows:

$$G = V^T x \qquad (G \in \{-1, +1\}) \quad \ldots\ldots \quad (1)$$

where, $V$ is the linear transformation matrix, $x$ is the input and $G$ is the transformed data in the LDA space.

For a two-class problem, the within-class scatter matrix $s_w$ consists of the sum of each within-class scatter matrix which means $s_{w0}$ and $s_{w1}$.The between-class scatter matrix is denoted with $s_b$. And those scatter matrices are defined as follows:

$$\mu_i = \frac{1}{N_i} \sum_{x \in X_i} x \tag{2}$$

$$s_{wi} = \sum_{x \in X_i} (x - \mu_i)(x - \mu_i)^T \tag{3}$$

$$s_w = s_{w0} + s_{w1} \tag{4}$$

$$s_b = (\mu_0 - \mu_1)(\mu_0 - \mu_1)^T \tag{5}$$

where the subscript $i$ indicates the category of the class and $\mu_i$ is the mean vector of the category of $X_i$. Because LDA demands to make the distance between the category centers of different types of data as large as possible, that is to maximize $s_b$. $N_i$ denotes the number of samples in $X_i$. At the same time, LDA hopes that the projection points of the unified category data are as close as possible. In other words, the covariance of the projection points of the same sample should be as small as possible after $s_w$ minimized. Thus, the above process can be considered as an optimization problem. According to [6,34], the optimization equation is described as follows :

$$\underbrace{arg\ max}_{V} J(V) = \frac{||V^T \mu_0 - V^T \mu_1||_2^2}{s_{w0}^2 + s_{w1}^2} \quad \ldots\ldots \quad (6)$$

However, equation (6) can be rewritten in the Ruili entropy form, which is $J(V) = \frac{V^T s_b V}{V^T s_w V}$. Hence, according to the properties of generalized Ruili entropy, the maximal value of $J$ is equal to the max eigenvalue of matrix $s_w^{-1} s_b$. Then, $V$ is the eigenvector corresponding to the $s_w^{-1} s_b$ largest eigenvalues $\lambda$. As LDA applied in two classification case, $s_b V$ is always in the same direction as $\mu_0 - \mu_1$. Let $s_b V = \lambda(\mu_0 - \mu_1)$, and substitute it in $(s_w^{-1} s_b)V = \lambda V$. Finally, solving the equation gives $V = s_w^{-1}(\mu_0 - \mu_1)$.

### 2.3. Ensemble Learning

Ensemble learning is a strategy for improving the model predictions of any given learning algorithm. The idea of this strategy is to train weak classifiers sequentially, each trying to compensate its predecessor. AdaBoost (Adaptive boosting) is a specific ensemble learning algorithm developed for classification problems which combines multiple weak classifiers into a single boosted classifier. The weak classifier is denoted as C, and its classifier error is defined as follows :

$$\text{error} = \frac{\sum_{j=1}^{n} w_j \mathrm{I}(\mathrm{C}(x_j) \neq y_j)}{\sum_{j=1}^{n} w_j} \tag{7}$$

$$\mathrm{I}(x) = \begin{cases} 0, & \text{if} \quad \mathrm{C}(x_j) = y_j \\ 1, & \text{if} \quad \mathrm{C}(x_j) \neq y_j \end{cases} \tag{8}$$

where $y$ is the label of the features vector with sample weights $w$. $\mathrm{I}(x)$ is a discriminant function whose output depends on whether the input variables are equal. AdaBoost trains a sequence of models with augmented sample weights, generating coefficients $\alpha$ for individual classifiers based on errors. Low errors lead to large $\alpha$, which means higher importance in final result voting. For more detail, in each iteration, AdaBoost identifies missclassified data points, increasing their weights and decrease the weights of correct points, such that the next classifier will pay extra attention to get them right. This process is repeated until the classification error meets the requirements. Finally, the output boosted classifier is to combine the sequence of models with $\alpha$.

In this study, we proposed the diabetes type classification algorithm based on LDA and AdaBoost. The details are provided in Algorithm 1.

## 3. Results

This section evaluates the performance of the proposed diabetes classification algorithm through an application of FGM data obtained from real diabetes participants.

### 3.1. Data Collection

In this study, the dataset was collected from 113 diabetes participants who wore FGM. The FGM system (Freestyle Libre H, Abbott, US) included an electrochemical sensor based on glucose oxidase, which was placed subcutaneously. The sensor could normally work for 14

**Algorithm 1** Ensemble-based LDA Classification Algorithm

**Input:**

$x, y$ : FGM training set $\{(x_1, y_1), (x_2, y_2) \ldots (x_n, y_n)\}$

$w$ : Initializing the weight distribution value of the training data, $w_0 = \frac{1}{n}$

C : Practicing LDA as a weak classifier denoted by C

$M$ : The number of iterations

**Output:**

$G$ : Final boosted classifier

1: **Initialize:** $w$, C

2: **while** $m < M$ **do**

3: **Classifier Training** using $w$ to generate weighted data samples. According to equation (1)-(6) fit the classifier C and output the classifier $\mathrm{C}^m$:

$$\mathrm{C}^m(x) = V^T x \quad (9)$$

4: **Calculate:**

$$\mathrm{error}^m = \frac{\sum_{j=1}^{n} w_m \mathrm{I}(\mathrm{C}^m(x_j) \neq y_j)}{\sum_{j=1}^{n} w_m} \quad (10)$$

$$\alpha^m = \frac{1}{2} \log \frac{1 - \mathrm{error}^m}{\mathrm{error}^m} \quad (11)$$

5: **Update:**

Update the weight distribution of the training dataset

$$w_m = w_m * e^{\alpha^m * \mathrm{I}(\mathrm{C}^m(x_j) \neq y_j)} \quad (12)$$

6: **Weights Normalized:**

$$w = w - mean(w) \quad (13)$$

7: **end while**

8: Building a linear combination of basic classifiers

$$f(x) = \sum_{m=1}^{M} \alpha_m * \mathrm{C}^m(x) \quad (14)$$

9: **Output Classification Result :**

$$G(x) = sign(f(x)) \quad (15)$$

10: **return** $G$

days, with a receiver that wirelessly collects stored interstitial glucose data in per 15 minutes. All of the participants were consecutively recruited from January 2018 to June 2019 who were hospitalized patients with diabetes at the Department of Endocrinology and Metabolism of Peking University People's Hospital. The FGM dataset consisted of a time-sequence flash glucose data. For each participant the FGM data recorded length was unequal, varying from a week to several months. Types of diabetes were diagnosed according to the 1999 World Health Organization (WHO) criteria. Eligibility criteria at recruiting were 18-75 years old with an established diabetes diagnosis and clear classification. The diabetes classification were made by an endocrinologist and independently confirmed by another specialist. Exclusion criteria included current use of non-FGM continuous glucose monitoring system, renal or liver failure, the inability of using the FGM system, the known allergy to medical-grade adhesives, and pregnancy. The study was approved by the institutional review board at Peking University People's Hospital and received written informed consent of all participants. As a result, after the data preprocessing, a total of 1,445 independent samples were generated to evaluate the model and divided into 2 parts, 90% of which were used for training and 10% were used for testing.

### 3.2. Raw Data Process and Visualization

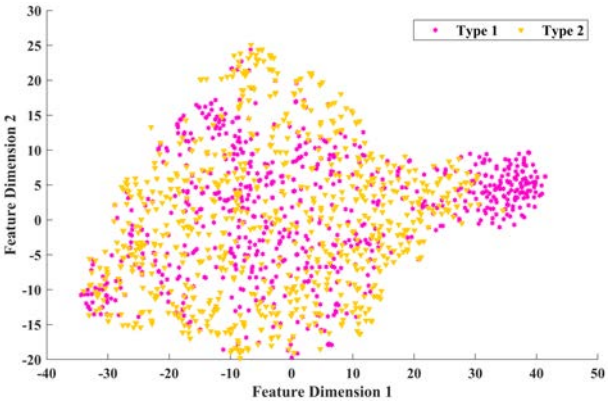


**Fig. 2.** The original dataset visualization. It can be seen from the figure that the samples of type 1 and type 2 are very evenly distributed, and there is no obvious decision surface.

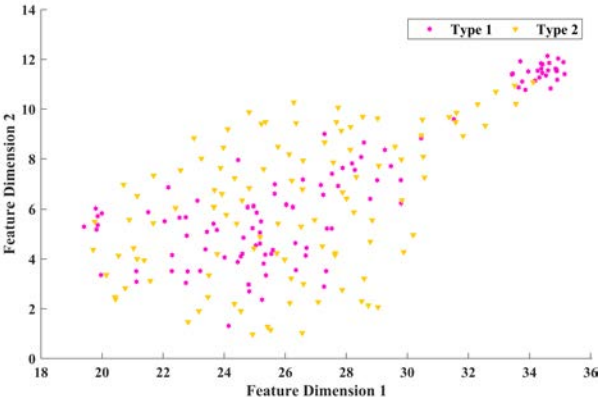


**Fig. 3.** Data visualization which has been down-sampling and regrouped. It can be seen from the figure that after the data preprocessing, the distribution of type 1 and type 2 samples has relatively clear cluster centers, and in particular, type 1 gradually converges toward the center. At the same time, due to the existence of downsampling, the distribution has obvious sparsity after feature compression.

In the data preprocessing stage, we downsample the FGM data and reorganize it, which effectively improves the accuracy of classification. To better analyze the data samples before and after preprocessing, we utilize the t-SNE method [31, 33] in this work to visualize the data samples, which is shown in Fig. 2 and Fig. 3. The method of t-SNE is generally used for data dimensionality reduction. In our experiments, we employ the t-SNE method to project FGM data from 96-dimensional to 2-bit space for realizing data visualization.

Fig. 2 shows the visualization of the original data set. There is no obvious dividing line of positive and negative

sample points. So that, it is difficult to achieve the classification of positive and negative samples by linear segmentation. Fig. 2 shows the visualization effect after the data is down-sampled and recombined. As a whole, the distribution of samples after down-sampling has a more sparse visualization effect and the distance between classes of type 1 and type 2 sample points increases. In the meantime, type 1 sample points are generally distributed on a progressive straight line.

### 3.3. Classification Performance

To evaluate the performance of our proposed algorithm, we employ a 10-fold cross-validated mean and some standard measures, such as specificity (SPE), precision (PRE), sensitivity (SEN), accuracy (ACC), F1-score (F1) and Matthews correlation coefficient (MCC). Ten times cross-validation scheme with randomization is employed to the dataset from each subject in order to increase the reliability of the classification results. In 10-fold cross-validation, the whole FGM dataset is divided into ten subsets. The proposed algorithm is trained with nine subsets of data sample vectors, whereas the remaining subset is used for testing. This procedure is repeated ten times as each subset has an equal chance of being the testing data. The average classification performances compared with other published classical literatures are reported in Table 1. From Table 1, the indicator difference is positive number, which shows that ensemble model can improve the performance of single model. But, the SPE, PRE difference in KNN and the PRE, SEN, F1, MCC difference in SVM decreased, which is marked in red. In addition, the decrease is discussed in discussion. In that case, ensemble learning does not improve significantly on these two methods. The highest accuracy of the single model is SVM algorithm, which is 68.26%, and the accuracy of the ensemble model is 69.70%, which performs best in the existing test methods.

From Table 2 the proposed methods is at the same level of the best performance among the references. The best classification accuracy of 87.88% is obtained from our proposed method. In the experimental results, the differences between the SEN and SPE indicators of the algorithms are slightly larger. Our proposed algorithm has the smallest difference between these two indicators (SPE is 0.9091 and SEN is 0.8182). Therefore, it has the highest score on the F1 indicator, which is 0.8182. That reflecting our method has a good classification ability for both positive and negative samples. Meanwhile, it can be known from Table 2 that our proposed algorithm performs best in the Matthews correlation coefficient (MCC score is 0.7273). It can be seen that our algorithm has better robustness and reliability than our compared algorithms in the classification of diabetes for blood glucose data.

## 4. Discussion

This study demonstrates the feasibility of using FGM data to recognize diabetes type. Several important issues are explored. Firstly, our work validates that the ensemble model can effectively improve single-model classification algorithms performance in diabetes classification problems based on FGM data. As shown in Fig. 4, the ensemble learning algorithm improves the performance of the single LR model by 74.82%. Comparing with other single models, the LR algorithm performance significantly improves after practicing the ensemble algorithm. In contrast, the performance improvement of the SVM algorithm after practicing the ensemble algorithm is not obvious, only improves of 0.0401%. The SVM and LR models are the typical representatives of strong classifiers and weak classifiers, respectively. From the F1 score in Table 1, other weak classifiers have been improved, but the indicators of the SVM algorithm decreased. Therefore, it is not difficult to find that the application effect of the ensemble algorithm is more inclined to enhance the weak classifier performance, which is consistent with the original intention of the ensemble algorithm. Comparing with the traditional ensemble learning algorithm, our proposed algorithm can solve this problem well, namely, the classification ability of the classifier is weakened by the recombination of the data, so that the ensemble algorithm takes effect. Comparing with previous methods that require a large amount of clinical data to obtain a diagnosis of diabetes typing, an ensemble model based on FGM data can effectively improve the efficiency of classification diagnosis.

Secondly, for the classification of sequence data, the traditional model structure generally consists of three parts: data preprocessing, feature extraction, and classifier. The effectiveness of traditional model classification depends on the extraction of data features and the selection of features. However, FGM data is composed of a set of dynamically changing FGM sequence numbers. As for diabetes typing, FGM data does not have notable waveform discrimination criteria similar to ECG data. It is difficult to select clinically interpretable feature types. The biological rhythm of the human body generally takes one day as a cycle, and the fluctuation of FGM values within a day affected by the interference factors such as diet and exercise. Therefore, we explore the method of fragment data reorganization with the cycle of biological rhythm. By converting the historical FGM data of the same patient into a combination of fragments from one or several biorhythmic cycles, a new sample of data recombination is constructed. This simplifies the link of explicit feature extraction. From the comparison of Tables 1 and 2, Fig. 2 and Fig. 3, it can be seen that the new data sample has a larger inter-class variance and a stronger ability to represent features.

Furthermore, from the observed effects before and after data processing, the down-sampling and reorganization of FGM data changes the distribution of the raw data, especially the time information. So that, the new sample could compress more time information and highlight the dependent connections between adjacent time points. Fig. 5 shows that down-sampling and reorganizing the original FGM data can effectively improve the classification per-

**Table 1.** Performance comparison between single model and single model plus combination learning

| | | ACC | SPE | PRE | SEN | F1 | MCC |
|---|---|---|---|---|---|---|---|
| LDA | Single | 0.4671 | 0.5096 | 0.3289 | 0.3968 | 0.3597 | -0.0911 |
| | Boosting | 0.6587 | 0.6992 | 0.4365 | 0.5612 | 0.4911 | 0.2446 |
| | **Difference** | 0.1916 | 0.1896 | 0.1076 | 0.1644 | 0.1314 | 0.3357 |
| KNN | Single | 0.6377 | 0.9087 | 0.5581 | 0.1905 | 0.284 | 0.1435 |
| | Boosting | 0.6557 | 0.6809 | 0.3492 | 0.5714 | 0.4335 | 0.2193 |
| | **Difference** | 0.018 | -0.2278 | -0.2089 | 0.3809 | 0.1495 | 0.0758 |
| NB | Single | 0.3563 | 0.3077 | 0.2764 | 0.4365 | 0.3385 | -0.2527 |
| | Boosting | 0.4012 | 0.5357 | 0.5873 | 0.3333 | 0.4253 | -0.1276 |
| | **Difference** | 0.0449 | 0.228 | 0.3109 | -0.1032 | 0.0868 | 0.1251 |
| LR | Single | 0.3772 | 0.6058 | 0 | 0 | Nan | -0.444 |
| | Boosting | 0.6557 | 0.6996 | 0.4444 | 0.5545 | 0.4934 | -0.2407 |
| | **Difference** | 0.2785 | 0.0938 | 0.4444 | 0.5545 | VALUE | 0.2033 |
| SVM | Single | 0.6826 | 0.7404 | 0.5781 | 0.5873 | 0.5827 | 0.3267 |
| | Boosting | 0.697 | 0.7727 | 0.5455 | 0.5455 | 0.5455 | 0.3182 |
| | **Difference** | 0.0144 | 0.0323 | -0.032 | -0.0418 | -0.0372 | -0.0085 |

**Table 2.** Ensemble Classifier with Data preprocessing Strategy

| | **Proposed** | KNN | NB | LR | SVM |
|---|---|---|---|---|---|
| ACC | **0.8788** | 0.7273 | 0.4848 | 0.8182 | 0.8485 |
| SPE | **0.9091** | 1 | 0.7778 | 0.9444 | 0.8636 |
| PRE | **0.8182** | 1 | 0.8182 | 0.9091 | 0.7500 |
| SEN | **0.8182** | 0.5500 | 0.3750 | 0.6667 | 0.8182 |
| F1 | **0.8182** | 0.7097 | 0.5143 | 0.7692 | 0.7826 |
| MCC | **0.7273** | 0.5701 | 0.1443 | 0.6455 | 0.6682 |

formance of the classification model.

In addition, we experimentally verify previous studies of weak classifier selection criteria in ensemble learning, namely, the closer the single model classification accuracy is to 50%, the better the ensemble learning effect is. As shown in Table 1, the single model classification accuracy of LDA is 46.71%, which is the closest to 50% compared with the single model classification accuracy of other methods.

## 5. Conclusion

In this work, we have proposed an ensemble-based LDA classification algorithm which has an accuracy of 87.88% and specificity of 90.91%. The novelty of this work lies in the data preprocessing and the combination between ensemble and LDA methods for diabetes classification. In addition, previous research has focused on diabetes prediction and classification of a single type of diabetes. Therefore, we have implemented a method that can classify two types of diabetes simultaneously. However,

the number of ensemble trainings and the choice of weak classifiers for classification methods based on ensemble learning directly affect the performance of classification. Although too many weak classifiers are integrated, the accuracy of the classification is improved, but at the same time it will cause over fitting problems. Therefore, our future research goals will pay more attention to solving the parameter adaptation problem in ensemble learning and the balancing method of classification models between over fitting and under fitting.

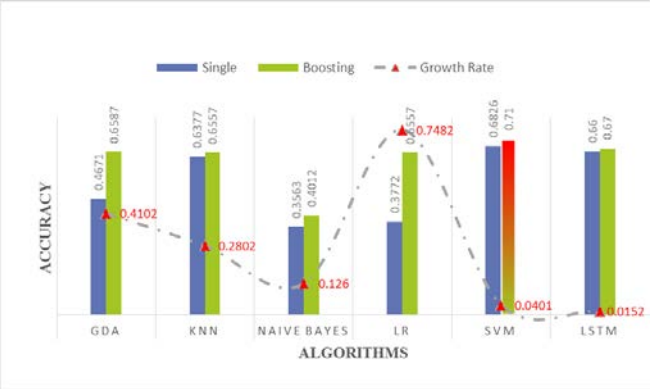


**Fig. 4.** Algorithm performance changes before and after ensemble learning employed.

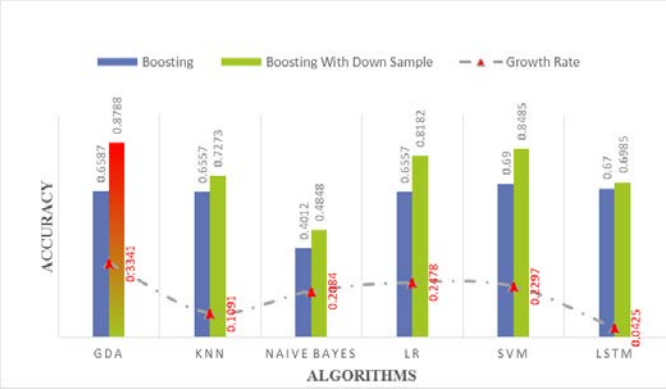


**Fig. 5.** When LDA is selected as the weak classifier in our proposed model, the performance is optimal.

**Acknowledgements**

The authors thank all the study participants, research staff, physicians and nurses in Peking University People's Hospital and students who participated in this work.

**References:**

[1] "Diagnosis and Classification of Diabetes Mellitus", Diabetes Care, 37(Supplement 1):S81–S90, 2014.

[2] K. M. Adal, P. G. van Etten, J. P. Martinez, K. W. Rouwen, K. A. Vermeer, and L. J. van Vliet, "An Automated System for the Detection and Classification of Retinal Changes Due to Red Lesions in Longitudinal Fundus Images", IEEE Transactions on Biomedical Engineering, 65(6):1382–1390, June 2018.

[3] W. AZ, r. E, B. T, A. D, M. L, and H. G, "Data-Driven Blood Glucose Pattern Classification and Anomalies Detection: Machine-Learning Applications in Type 1 Diabetes", J Med Internet Res, 21(5):e11030, 2019.

[4] J. Carter, C. Long, B. Smith, T. Smith, and T. Donati, "Combining elemental analysis of toenails and machine learning techniques as a non-invasive diagnostic tool for the robust classification of type-2 diabetes", Expert Systems with Applications, 115:245 – 255, 2019.

[5] J. Casqueiro and C. Alves, "Infections in patients with diabetes mellitus: a review of pathogenesis", IndianJournal of Endocrinology and Metabolism, 16(7):22, 2012.

[6] C.-K. Chen and J.-Y. Yang, "Two-Dimensional Discriminant Transform Based on Scatter Difference Criterion for Face Recognition", pp. 683–686, 09 2006.

[7] C. Davi, A. Pastor, T. Oliveira, F. B. d. L. Neto, U. Braga-Neto, A. W. Bigham, M. Bamshad, E. T. A. Marques, and B. Acioli-Santos, "Severe Dengue Prognosis Using Human Genome Data and Machine Learning", IEEE Transactions on Biomedical Engineering, 66(10):2861–2868, Oct 2019.

[8] A. Dinh, S. Mietschin, A. Young, and S. Mohanty, "A data-driven approach to predicting diabetes and cardiovascular disease with machine learning", BMC Med Inform Decis Mak, 6:211, 2019.

[9] Dworzynski, P., M. Aasbrenn, Rostgaard, and K. e. al., "Nationwide prediction of type 2 diabetes comorbidities", Sci Rep, 10:1776, 6 2020.

[10] B. I. Frohnert, B.-J. Webb-Robertson, L. M. Bramer, S. M. Reehl, K. Waugh, A. K. Steck, J. M. Norris, and M. Rewers, "Predictive Modeling of Type 1 Diabetes Stages Using Disparate Data Sources", Diabetes, 69(2):238–248, 2020.

[11] H. Fröhlich, K. Claes, C. De Wolf, X. Van Damme, and A. Michel, "A Machine Learning Approach to Automated Gait Analysis for the Noldus Catwalk System", IEEE Transactions on Biomedical Engineering, 65(5):1133–1139, May 2018.

[12] Hathaway, Q.A, Roth, S.M, Pinti, M.V, and e. al, "Machine-learning to stratify diabetic patients using novel cardiac biomarkers and integrative genomics", Cardiovasc Diabetol, 18:78, 2019.

[13] F. Heslinga, J. Pluim, B. A. Houben, M. Schram, R. Henry, C. Stehouwer, M. v. Greevenbroek, T. T. Berendschot, and M. Veta, "Direct Classification of Type 2 Diabetes From Retinal Fundus Images in a Population-based Sample From The Maastricht Study", Preprint, 11 2019.

[14] K. Inoue, M. Yoshioka, N. Yagi, S. Nagami, and Y. Oku, "Using Machine Learning and a Combination of Respiratory Flow, Laryngeal Motion, and Swallowing Sounds to Classify Safe and Unsafe Swallowing", IEEE Transactions on Biomedical Engineering, 65(11):2529–2541, Nov 2018.

[15] D. Jacobs, T. Hilton, M. del Campo, P. L. Carlen, and B. L. Bardakjian, "Classification of Pre-Clinical Seizure States Using Scalp EEG Cross-Frequency Coupling Features", IEEE Transactions on Biomedical Engineering, 65(11):2440–2449, Nov 2018.

[16] T. Kamali and D. W. Stashuk, "Electrophysiological Muscle Classification Using Multiple Instance Learning and Unsupervised Time and Spectral Domain Analysis", IEEE Transactions on Biomedical Engineering, 65(11):2494–2502, Nov 2018.

[17] K. Kannadasan, D. R. Edla, and V. Kuppili, "Type 2 diabetes data classification using stacked autoencoders in deep neural networks", Clinical Epidemiology and Global Health, 7(4):530 – 535, 2019.

[18] K. Kannadasan, D. R. Edla, and V. Kuppili, "Classification and prediction of diabetes disease using machine learning paradigm", Health Inf Sci Syst, 8, 7 2020.

[19] B. Lei, P. Yang, Y. Zhuo, F. Zhou, D. Ni, S. Chen, X. Xiao, and T. Wang, "Neuroimaging Retrieval via Adaptive Ensemble Manifold Learning for Brain Disease Diagnosis", IEEE Journal of Biomedical and Health Informatics, 23(4):1661–1673, July 2019.

[20] R. K. Maji, S. Khatua, and Z. Ghosh, "A Supervised Ensemble Approach for Sensitive microRNA Target Prediction", IEEE/ACM Transactions on Computational Biology and Bioinformatics, 17(1):37–46, Jan 2020.

[21] S. Mirzaei, T. Sidi, C. Keasar, and S. Crivelli, "Purely Structural Protein Scoring Functions Using Support Vector Machine and Ensemble Learning", IEEE/ACM Transactions on Computational Biology and Bioinformatics, 16(5):1515–1523, Sep. 2019.

[22] K. Ogurtsova, J. Rocha Fernandes, Y. Huang, U. Linnenkamp, L. Guariguata, N. Cho, D. Cavan, J. Shaw, and L. Makaroff, "IDF diabetes atlas: globalestimates for the prevalence of diabetes for 2015 and 2040", Diabetes Res Clin Pract, 128(jun):40–50, 2017.

[23] J. Prince, F. Andreotti, and M. De Vos, "Multi-Source Ensemble Learning for the Remote Prediction of Parkinson's Disease in the Presence of Source-Wise Missing Data", IEEE Transactions on Biomedical Engineering, 66(5):1402–1411, May 2019.

[24] Qiu, H., Yu, H., Wang, L., and e. al, "Electronic Health Record Driven Prediction for Gestational Diabetes Mellitus in Early Pregnancy", Sci Rep, 7:16417, 2017.

[25] M. N. I. Qureshi, B. Min, H. Park, D. Cho, W. Choi, and B. Lee, "Multiclass Classification of Word Imagination Speech With Hybrid Connectivity Features", IEEE Transactions on Biomedical Engineering, 65(10):2168–2177, Oct 2018.

[26] S. Raghu, N. Sriraam, G. P. Kumar, and A. S. Hegde, "A Novel Approach for Real-Time Recognition of Epileptic Seizures Using Minimum Variance Modified Fuzzy Entropy", IEEE Transactions on Biomedical Engineering, 65(11):2612–2621, Nov 2018.

[27] R. V. Sharan, U. R. Abeyratne, V. R. Swarnkar, and P. Porter, "Automatic Croup Diagnosis Using Cough Sound Recognition", IEEE Transactions on Biomedical Engineering, 66(2):485–495, Feb 2019.

[28] R. Tao, S. Zhang, X. Huang, M. Tao, J. Ma, S. Ma, C. Zhang, T. Zhang, F. Tang, J. Lu, C. Shen, and X. Xie, "Magnetocardiography-Based Ischemic Heart Disease Detection and Localization Using Machine Learning Methods", IEEE Transactions on Biomedical Engineering, 66(6):1658–1667, June 2019.

[29] Y. Tao, Y. J. Chen, X. Fu, B. Jiang, and Y. Zhang, "Evolutionary Ensemble Learning Algorithm to Modeling of Warfarin Dose Prediction for Chinese", IEEE Journal of Biomedical and Health Informatics, 23(1):395–406, Jan 2019.

[30] Y. Tao, Y. J. Chen, L. Xue, C. Xie, B. Jiang, and Y. Zhang, "An Ensemble Model With Clustering Assumption for Warfarin Dose Prediction in Chinese Patients", IEEE Journal of Biomedical and Health Informatics, 23(6):2642–2654, Nov 2019.

[31] L. v. d. Maaten, "Barnes-Hut-SNE", CoRR, abs/1301.3342, 2013.

[32] K. Varma and B. Panda, "Comparative analysis of Predicting Diabetes Using Machine Learning Techniques", 6:522–530, 6 2019.

[33] M. Wattenberg, F. Viégas, and I. Johnson, "How to Use t-SNE Effectively", Computer Science, 2016.

[34] J. Ye, R. Janardan, and Q. Li, "Two-Dimensional Linear Discriminant Analysis.", Adv. Neural. Inf. Process. Syst., 17, 01 2004.