Indonesian Journal of Electrical Engineering and Computer Science
Vol. 36, No. 1, October 2024, pp. 672~684
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v36.i1.pp672-684

Experimental study of a medical data analysis model based on comparative performance of classification algorithms

Aigerim Ismukhamedova[1], Indira Uvaliyeva[2], Zhenisgul Rakhmetullina[2]
[1]School of Computer Science, D. Serikbayev East Kazakhstan Technical University, Oskemen, Kazakhstan
[2]School of Digital Technology and Intelligent Systems, D. Serikbayev East Kazakhstan Technical University, Oskemen, Kazakhstan

Article Info

Article history:

Received Mar 28, 2024
Revised Jun 5, 2024
Accepted Jun 25, 2024

Keywords:

Diabetes prevalence
Digital healthcare
Electronic health system
Health informatics
Healthcare resource allocation
Intelligent decision support
Medical data research
Support of medical decisions

ABSTRACT

This article centers around the development and analysis of machine learning (ML) and deep learning models aimed at enhancing diabetes diagnosis. In the swiftly evolving landscape of data technologies, it becomes crucial to explore the applications of these methods for accurate predictions and improved medical decision-making. Our research encompasses diverse datasets, leveraging state-of-the-art algorithms and technologies for model training and testing. The primary emphasis lies in evaluating the accuracy, sensitivity, and specificity of models within the realm of diabetes diagnosis. The study results reveal significant advancements in disease prediction, underscoring the potential of ML and deep learning in medical applications. This work introduces fresh perspectives on the utilization of computational methods in healthcare and serves as a foundation for prospective research in this domain.

Corresponding Author:

Indira Uvaliyeva
School of Digital Technology and Intelligent Systems
D. Serikbayev East Kazakhstan Technical University
Oskemen, East Kazakhstan, Kazakhstan
Email: indirauvalieva@gmail.com

1. INTRODUCTION

In the evolving landscape of modern healthcare, the management and analysis of health data have become pivotal. The introduction of the electronic health passport (EHP) represents a leap towards enhancing patient data handling, offering a comprehensive tool for the collection, storage, and processing of health information. This innovation aligns with the World Health Organization's (WHO) Global Strategy for Digital Health 2020-2025 [1], [2] which aspires to universally improve health care via digital technologies [3], [4] with an emphasis on equity [5] and inclusion [6]. Despite these advances, the implementation of such strategies across varying national landscapes poses considerable challenges.

Background: The foundation for constructing systems that gather, store, and analyze medical information from patients across various countries globally lies in the realm of extensive data. Processing this vast amount of data, commonly referred to as big data, empowers us to formulate methodologies for anticipating factors such as illness rates, mortality, complications, and beyond [7]. Big data processing has made it possible to develop intelligent decision-making systems, including in medicine [8]–[12]. Authors contend that the application of artificial intelligence (AI) and machine learning (ML) has contributed to enhancing outcomes in the diagnosis, treatment, and prognosis for Chronic Limb-Threatening Ischemia (CLTI) patients [13].

The efficacy of employing ML technologies in medicine is substantiated by numerous studies across various medical fields. In their examination of the diagnostic properties of ML algorithms in peripheral artery disease, the authors conclude that ML enables more precise classification and prediction of the disease [14].

ML methodologies are employed to predict biological age by utilizing data linked to identifiable mental traits that are correlated with accelerated aging [15]. The authors of studies on Parkinson's disease [16], high-fatality cancers [17]–[20], and inherited arrhythmia syndromes [21] reach the same conclusion. The utilization of deep learning algorithms reduces the likelihood of a false-positive diagnosis, thereby eliminating the negative psychological impact faced by the patient [22]. The utilization of ML methods and deep learning algorithms in diagnosing COVID-19 contributed to bringing the pandemic under control [23]-[35].

ML and deep learning methods, with their inherent capacity to independently extract valuable insights from data, offer a significant advantage in prediction, leading to their widespread utilization in the medical field, particularly underscoring the importance and high significance of DL Methods in healthcare [36]–[39]. Analyzing patient records and selection criteria for developed models enables predictions of contagion, survivability, and the risk of critical conditions [11], [40]–[43]. A comparison of 14 different ML algorithms revealed an accuracy of over 90% in identifying individuals with and without diabetes [44]. Researchers combine multiple algorithms to determine the accuracy of predicting diabetes [45]–[47].

Researchers have demonstrated the efficacy of ML models in disease identification and prognosis with notable accuracy. Yet, the utility of these technologies in the nuanced field of diabetes prediction remains underexplored. This occurs partly due to the algorithms employed. For instance, within the realm of alternative ML algorithms, conventional methods like naive Bayes, logistic regression, and support vector machines (SVMs) lead to exponential growth in computational complexity due to data expansion, resulting in inadequate outcomes [7], [33], [48], [49]. In contrast, tree algorithms provide a more robust alternative, alleviating certain limitations encountered by traditional methods [8], [9], [50]. Ensemble methods and the decision tree algorithm in ML offer approaches to high-precision diagnosis and prognosis for breast cancer [51]-[53], cardiovascular disease [54], [55], and COVID-19 [56].

The problem: The widespread application of ML across the nexus of medicine and information technology has significantly advanced our capabilities to address various challenges. Despite these advancements, the quest for innovative diagnostic methodologies through the meticulous analysis of patient data remains largely unexplored. Prevailing literature underscores the incremental adoption of digital health technologies, with strategic objectives focusing on governance, resource capacity, and data security. However, literature reveals a gap in the pragmatic analysis of big data to inform healthcare delivery. Existing research predominantly relies on established disease markers or singular types of medical information.

The proposed solution: In contrast, we advocate for a novel methodology that leverages ML to uncover new diagnostic indicators. This approach aims to elucidate the connections between diseases and a comprehensive array of patient medical data, encompassing prior illnesses and all available health information. Such a strategy is not only pivotal but also necessitated by the integration of cutting-edge technologies within routine medical frameworks and the adoption of electronic health passports. We will elucidate our methodology, combining various ML models with deep learning classification techniques to forecast diabetes presence. The research will undertake a binary classification task, employing multiple evaluation criteria and hyperparameter tuning to forge robust algorithms. Through meticulous data preprocessing, we will also explore patient history and hospitalization levels to enhance our model's predictive capability.

2. MATERIALS AND METHODS

2.1. Research description

The study progresses through methodical steps designed to ensure replicability and robustness of the findings. Below is a detailed breakdown of each phase.

A. Selection and preparation of data:
- Exploratory data analysis (EDA) is conducted to identify and address data inconsistencies, missing values, and outliers.
- Data quality is reassessed through verification protocols to ensure validity and reliability.
- Structured dataframes are constructed for testing and training predictive models, with careful consideration of feature selection to balance the informative value against computational efficiency.

B. Modeling:
- A range of predictive models are delineated, including but not limited to, decision trees, random forests, and SVMs, each chosen for their proven efficacy in medical data analysis [38], [39], [56], [57].
- Model performance is rigorously evaluated against key metrics: accuracy, F1 score, precision, and recall to ensure a comprehensive understanding of each model's predictive power.

C. Identification of the best model:
- Comparative analysis is conducted to determine the most effective model. This involves a systematic review of performance metrics and alignment with the specific nuances of the dataset.

D. Optimization of the model:
- Model accuracy is enhanced through a systematic optimization process utilizing Grid Search techniques to fine-tune hyperparameters within computationally feasible bounds.

E. Application of deep learning:
- Ensemble models and neural networks are deployed, selected for their capacity to model complex non-linear relationships inherent in medical data [8], [9], [52]–[58].
- Hyperparameters are carefully chosen based on their impact on performance, with a focus on generalizability and prevention of overfitting.

F. Obtaining and analyzing results:
- Results are acquired from a randomized dataset to ensure the robustness of findings.
- Key parameters influencing diagnosis are identified through feature importance analysis, which informs the interpretability of the model.
- Collaboration with domain experts ensures the clinical relevance of the data analysis.
- Additional testing is conducted where possible, using datasets from medical institutions to validate the model's applicability in a real-world setting.

Data selection: Data selection for our study on diabetes mellitus follows diagnostic and treatment protocols, adhering to WHO standards, and utilizes indicators like symptoms and glucose levels for identification. We extracted and analyzed a comprehensive dataset from the MIMIC III database, covering various medical and patient-related categories to ensure data integrity. Data processing and model training: In data processing and model training, we used a broad range of indicators including laboratory and patient examination data. Our approach included data cleansing to eliminate duplicates and anomalies, followed by data integration to support our research objectives. Insights were drawn from specific MIMIC database tables like 'chartevents' and 'noteevents' to enhance our understanding of patient status and treatment efficacy.

Optimal model determination: Subsequent to data preparation, we employ a combination of ML and ensemble methods to identify the optimal model. The selection of these methods is justified by their robustness in handling high-dimensional data and their proven track record in medical data analysis.

2.2. ML and ensemble methods

Traditional ML models like Logistic Regression use a linear decision boundary, which may struggle with complex, high-dimensional data. To address this, ensemble methods such as Random Forests, which use multiple Decision Trees, and Bagging and Boosting techniques are used to improve model robustness and accuracy. Bagging trains multiple models on different data subsets and aggregates their predictions, while Boosting sequentially corrects the previous model's errors, weighing predictions based on their accuracy [57].

Artificial neural networks (ANNs), including convolutional neural networks (CNNs) and recurrent neural networks (RNNs) [59]-[61], use layers of neurons to process data. CNNs are effective for automatic feature extraction through convolutional filters [62]-[64], whereas RNNs handle sequential data, with outputs from one layer fed back into the same layer as input [65]–[67]. Long short-term memory networks (LSTMs) improve RNNs by managing information flow with gates that regulate input, output, and forget functions [68].

Key to training these networks is minimizing a Cost Function [69], often using a Gradient Descent algorithm, which adjusts parameters to reduce prediction error [70]. Mini-batch gradient descent enhances efficiency by using subsets of data, while the Adam optimizer refines parameter updates by adjusting the learning rate based on gradient estimates [71]–[74]. These methods ensure effective learning and convergence during model training [65]–[74].

2.3. Optimization process

The optimization process involved the following steps, implemented using GridSearchCV, and explained through mathematical formulations:
- Activation function and kernel initializer: The scaled exponential linear unit (SELU) function, defined as in (1):

$$\begin{cases} x & if\ x > 0 \\ \alpha e^{x} - \alpha & if\ x \leq 0 \end{cases} \quad (1)$$

along with the LeCun normal initialization, has been employed to initialize weights and biases. SELU facilitates self-normalization by ensuring a mean of 0 and variance of 1 across layers, which is particularly effective in deep networks prone to the vanishing/exploding gradient problem.
- Learning rate and optimizers: Through GridSearchCV, the learning rate η and optimizers like Nadam, a variant of Adam incorporating Nesterov momentum, were evaluated for optimal performance.

– Epochs and batch size: The number of epochs and the mini-batch size *m* were optimized to balance the trade-off between computational efficiency and model performance.

By employing GridSearchCV, each hyperparameter was systematically varied and the model's performance was evaluated, leading to the selection of the hyperparameters that yielded the best validation performance.

2.4. Evaluation metrics for classification models

In assessing the performance of classification models within ML and DL. Several key metrics are used: precision, recall [75], F1-score [75], [76], accuracy, and the confusion matrix. The confusion matrix is a crucial tool that categorizes predictions into four types (Figure 1): True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

	Predicted class		
Actual Class		Class = Yes	Class = No
	Class = Yes	True Positive	False Negative
	Class = No	False Positive	True Negative

Figure 1. Confusion matrix or contingency table

These quadrants enable us to compute the following metrics:

A. Accuracy: The overall correctness of the model, presented in (2):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

B. Precision: the ratio of correct positive predictions to all positive predictions, presented in (3):

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

C. Recall (Sensitivity): the ratio of actual positives correctly identified, presented in (4):

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

D. F1-score: a harmonic mean of precision and recall, presented in (5):

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (5)$$

To systematize these computations, one may define a vector y representing actual class labels and a vector $\hat{y}$ for predicted labels. These vectors serve as inputs to functions calculating each metric, effectively transforming the classification problem into a mathematical evaluation. It is presented in (6):

$$metric(y, \hat{y}) \rightarrow R \quad (6)$$

where the function "metric" could be any of the beforementioned metrics, and R represents the real numbers, indicating that each metric provides a real-valued outcome representing the model's performance.

Further in the study we will build a visual representation in the form of a mixing matrix using the Python library Matplotlib or Seaborn, which allows an immediate graphical interpretation of the classifier results. In practice, metrics are of paramount importance when fine-tuning classification models, especially in domains with unbalanced datasets, where accuracy alone may not reflect the true predictive power of the model. Thus, the selection of appropriate metrics is necessary to ensure that the performance of the model is truly indicative of its ability to generalize to new, as-yet-unacquired data.

2.5. Data preparation and pre-processing

Data from the MIMIC-III database was used, focusing on diabetic patient records including demographics, medical history, medications, and hospitalizations. The broad scope of the data, beyond just diabetes-related information, was considered crucial for identifying new predictive factors. Data cleansing involved removing records with missing essential information, using imputation for some missing values, or

excluding such records entirely. Numerical data was standardized, for instance using the Z-transform, and categorical data was encoded using methods like one-hot encoding to facilitate ML processing. These preprocessing steps ensured data quality and consistency, critical for the accuracy of ML models. Model selection was based on data characteristics relevant to diabetes classification tasks.

Decision tree, random forest, and AdaBoost models were used for their unique advantages: decision trees for interpretability, random forests for handling large datasets and resistance to overfitting, and AdaBoost for boosting the performance of weaker classifiers [59]. Hyperparameter tuning was carried out using Grid Search or Random Search to optimize model settings, focusing on parameters like tree number and depth for random forests, and iteration count and learning rate for AdaBoost. This tuning helps improve model performance and address overfitting or underfitting. Feature importance analysis, like the feature_importances_ attribute in random forests, was utilized to identify key predictors such as blood glucose level, BMI, age, and gender, aiding in model interpretation and focusing on critical data aspects.

3. RESULTS AND DISCUSSION

3.1. Methods of comparison and results obtained

The performance of the models was assessed using metrics such as accuracy, F1-score, recall, and precision. This enables comparison of the models both in overall terms and in specific aspects of classification.

The following steps were undertaken:

A. Imported all the necessary libraries:
- pandas
- numpy
- matplotlip
- seaborn
- sklearn

B. Loaded processed diabetes data from the Diseases table and anonymized patient data from the Patients table. We also merged these data by key columns. We obtained the final dataframe with the following dimensions:
- Number of Patients with Diabetes: (552336, 13)
- Number of Patients without Diabetes: (15969, 13)
- Number of Final balanced Dataframe: (31938, 13)

C. Next, we begin our modeling
- Importing the necessary libraries. The code starts by importing classes from the sklearn library used for cross-validation and various classification algorithms.
- Set the initial state of the random number generator. Random_state = 2 ensures reproducibility of results by setting the initial state for the random number generator.
- Creating a list of classifiers. Various classification algorithms are added to the list of classifiers, including decision tree classifier and ensemble methods such as AdaBoost classifier, random forest classifier, extra trees classifier, and gradient boosting classifier.
- Using cross-validation with five partitions (kfold = 5), the code evaluates each classifier on the training dataset X_train with y_train responses using the accuracy metric.
- For each classifier, the average accuracy and standard deviation of the cross-validation results are computed.
- A bar chart is created that shows the average accuracy (mean of the cross-validation scores) for each algorithm with the standard deviation of the errors. The bar chart has been illustrated in Figure 2.

The diagram shows that gradient boosting classifier has the highest average accuracy while decision tree classifier has the lowest. This suggests that ensemble learning based models generally outperform the simple decision tree classifier, probably due to their ability to generalize better and reduce overfitting.

The Figure 3 delineates a quantitative evaluation of five distinct classification algorithms. Performance metrics are aggregated from cross-validation procedures, which are instrumental in assessing the robustness of each model. The metrics considered are the mean accuracy (CrossValMeans) and the associated standard error (CrossValerrors), derived from the cross-validation scores.

Variables described:
- CrossValMeans: This column represents the mean accuracy score obtained from the cross-validation process. Accuracy is a common metric in model evaluation, indicating the ratio of correctly predicted observations to the total observations.

– CrossValerrors: This column indicates the standard error of the mean accuracy scores across the cross-validation folds. A lower value here suggests higher stability in the model's performance across different data subsets.

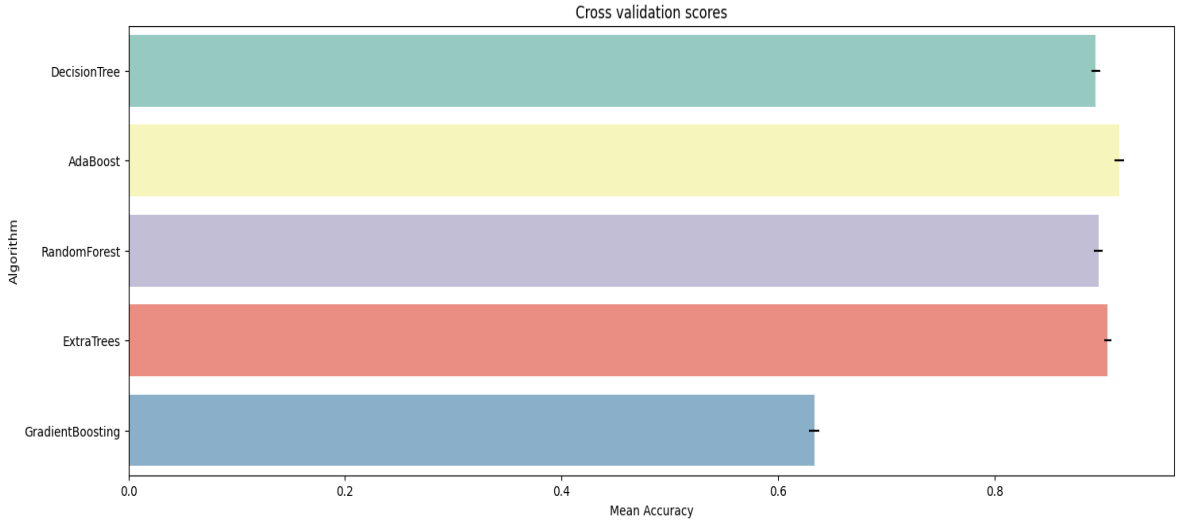

Figure 2. The diagram shows the cross-validation results for each of the algorithms

	CrossValMeans	CrossValerrors	Algorithm
0	0.893585	0.004231	DecisionTree
1	0.915459	0.004543	AdaBoost
2	0.896046	0.004023	RandomForest
3	0.904589	0.003482	ExtraTrees
4	0.633432	0.004944	GradientBoosting

Figure 3. Comparative performance of classification algorithms

From the data provided, we can infer the following:

– The AdaBoost algorithm exhibits the highest mean accuracy (0.915459) with a standard error of 0.004543, suggesting it not only performs well on average but also maintains consistent performance across different subsets of the data.
– The GradientBoosting algorithm shows significantly lower mean accuracy (0.633432) with a relatively higher standard error (0.004944), which could indicate less reliable performance across the cross-validation folds.
– The remaining algorithms — DecisionTree, RandomForest, and ExtraTrees — display moderate mean accuracies with DecisionTree having the least consistency, as suggested by the higher standard error.

This comparison is crucial for model selection in a data science project. The AdaBoost model would be a strong candidate for further refinement and testing, given its superior mean accuracy and consistent cross-validation performance. In contrast, the lower accuracy of GradientBoosting warrants an investigation into whether the model's complexity is appropriate for the data or if its parameters need adjustment.

D. Next, we started to implement hyperparameter selection for various ML algorithms using Grid Search parameter enumeration and estimation using cross-validation. Detailed description:

– Initialization of EstimatorSelectionHelper (EstimatorSelectionHelper). The constructor accepts dictionaries with models and their corresponding hyperparameters. It is checked that parameters are provided for each model.
– Fit function. A GridSearchCV is executed for each model, which systematically works through the combinations of parameters provided in params, evaluating each combination using cv fold cross-validation.
– Score_summary function. Collects GridSearchCV results and creates a summary table including the minimum, maximum, mean, and standard deviation of test scores. The table is sorted by the mean score (or other specified criterion) to show the best combinations of parameters.

- Creating models and params. In models and params, the classifiers and hyperparameters for the search are specified. For example, max_depth is selected for the DecisionTreeClassifier and learning_rate and n_estimators are selected for the AdaBoostClassifier.
- An instance of EstimatorSelectionHelper is created with the specified models and parameters.

The result is a console output of model names and a table that sorts the models by their performance given the configured hyperparameters. This table helps to determine which models and parameter sets work best for a given classification task. This table has been illustrated in Figure 4, which presents the results of hyperparameter selection for different classification algorithms using Grid Search. Each row of the table corresponds to a particular combination of parameters for a particular algorithm.

The table provides the following information on the hyperparameters:

- estimator. This parameter indicates the classifier for which the parameters were fitted.
- min_score, mean_score, max_score. These parameters indicate the minimum, mean and maximum accuracy achieved by the model on the cross-validation datasets.
- std_score. This parameter indicates the standard deviation of accuracy, reflecting the variation in model performance on different cross-validation folds.
- learning_rate, max_depth, max_features. These hyperparameters were picked up during the Grid Search process. NaN values mean that the parameter was not applied to this model.

The table shows that gradient boosting classifier generally performs better in terms of mean accuracy (mean_score) with different combinations of hyperparameters. Models with a maximum tree depth (max_depth) of 9 and max_features set to auto or sqrt give the best results. This may indicate that more complex models perform better, but it is important to keep an eye on overtraining. The learning_rate and max_features columns do not apply to all models due to the specific nature of the algorithms. For example, max_features does not apply to AdaBoost classifier and decision tree classifier in this setting.

	estimator	min_score	mean_score	max_score	std_score	learning_rate	max_depth	max_features
41	GradientBoostingClassifier	0.910952	0.914678	0.917058	0.00266821	NaN	9	auto
37	GradientBoostingClassifier	0.874507	0.876104	0.877888	0.00138689	NaN	7	auto
39	GradientBoostingClassifier	0.866804	0.870875	0.874319	0.0030996	NaN	9	sqrt
40	GradientBoostingClassifier	0.861732	0.866491	0.873943	0.00533642	NaN	9	auto
36	GradientBoostingClassifier	0.799361	0.809569	0.81796	0.00770064	NaN	7	auto
38	GradientBoostingClassifier	0.797577	0.802774	0.812324	0.00676151	NaN	9	sqrt
33	GradientBoostingClassifier	0.787808	0.793037	0.803119	0.00713072	NaN	5	auto
35	GradientBoostingClassifier	0.792316	0.79504	0.798892	0.00280016	NaN	7	sqrt

Figure 4. Results of models and selected hyperparameters

E. Subsequently, we conducted the final tuning and evaluation of several classification models on test data. Utilizing training and testing accuracy data, as well as precision, recall, and F1-score metrics for model comparison, the following procedures were executed:

- Importation of the metrics library. The code commences with the importation of functions for computing the model's quality metrics.
- Initialization of classifiers with specific hyperparameters. A list of classifiers is generated, each equipped with the optimal hyperparameters determined during the previous Grid Search phase.
- Model training and prediction. Each model is trained on the training dataset (X_train, y_train) and subsequently makes predictions on the test dataset (X_test).
- Calculation of metrics. For each model, accuracy, precision, recall, and F1-score are computed both on the training and the test datasets.
- Reporting of results. The calculated metrics for each model are outputted, along with a classification report and confusion matrix.

The calculated metrics, confusion matrix (7) and classification report as shown in Table 1 for DecisionTree model are represented below:

- Accuracy: 1.0; F1 score: 1.0; Recall: 1.0; Precision: 1.0

$$\begin{bmatrix} 3617 & 1165 \\ 2273 & 2527 \end{bmatrix} \quad (7)$$

Table 1. The classification report for decision tree model

	Precision	Recall	F1-score	Support
False	0.61	0.76	0.68	4782
True	0.68	0.53	0.60	4800
accuracy			0.64	9582
macro avg	0.65	0.64	0.64	9582
weighted avg	0.65	0.64	0.64	9582

The calculated metrics, confusion matrix (8) and classification report as shown in Table 2 for AdaBoost model are represented below:

– Accuracy: 1.0; F1 score: 1.0; Recall: 0.0; Precision: 1.0

$$\begin{bmatrix} 3526 & 1256 \\ 2720 & 2080 \end{bmatrix} \quad (8)$$

Table 2. The classification report for AdaBoost model

	Precision	Recall	F1-score	Support
False	0.56	0.74	0.64	4782
True	0.62	0.43	0.51	4800
accuracy			0.59	9582
macro avg	0.59	0.59	0.58	9582
weighted avg	0.59	0.59	0.58	9582

The calculated metrics, confusion matrix (9) and classification report as shown in Table 3 for RandomForest model are represented below:

– Accuracy: 1.0; F1 score: 1.0; Recall: 1.0; Precision: 1.0

$$\begin{bmatrix} 3511 & 1271 \\ 1877 & 2923 \end{bmatrix} \quad (9)$$

Table 3. The classification report for random forest model

	Precision	Recall	F1-score	Support
False	0.65	0.73	0.69	4782
True	0.70	0.61	0.65	4800
accuracy			0.67	9582
macro avg	0.67	0.67	0.67	9582
weighted avg	0.67	0.67	0.67	9582

The calculated metrics, confusion matrix (10) and classification report as shown in Table 4 for ExtraTrees model are represented below:

– Accuracy: 1.0; F1 score: 1.0; Recall: 1.0; Precision: 1.0

$$\begin{bmatrix} 3462 & 1320 \\ 2097 & 2703 \end{bmatrix} \quad (10)$$

Table 4. The classification report for ExtraTrees model

	Precision	Recall	F1-score	Support
False	0.62	0.72	0.67	4782
True	0.67	0.56	0.61	4800
accuracy			0.64	9582
macro avg	0.65	0.64	0.64	9582
weighted avg	0.65	0.64	0.64	9582

The calculated metrics, confusion matrix (11) and classification report as shown in Table 5 for GradientBoosting model are represented below:

– Accuracy: 1.0; F1 score: 1.0; Recall: 1.0; Precision: 1.0

$$\begin{bmatrix} 4655 & 127 \\ 675 & 4125 \end{bmatrix} \quad (11)$$

Table 5. The classification report for gradient boosting model

	Precision	Recall	F1-score	Support
False	0.87	0.97	0.92	4782
True	0.97	0.86	0.91	4800
accuracy			0.92	9582
macro avg	0.92	0.92	0.92	9582
weighted avg	0.92	0.92	0.92	9582

- Creating and outputting a summary table, represented in Figure 5. It collects information on each metric for all models.
- Graph construction. The subcategory bar function is used to construct a graph comparing the models on different metrics. The model comparison is represented in Figure 6.

Model	Train Accuracy	Test Accuracy	Precision	Recall	F1 Score
DecisionTree	0.658347	0.641202	0.684453	0.526458	0.595148
AdaBoost	0.590222	0.585055	0.623501	0.433333	0.511308
RandomForest	0.688853	0.671467	0.696948	0.608958	0.649989
ExtraTrees	0.658481	0.643394	0.671887	0.563125	0.612717
GradientBoosting	0.958356	0.916301	0.970132	0.859375	0.911401

Figure 5. Summary table of each metric for all models

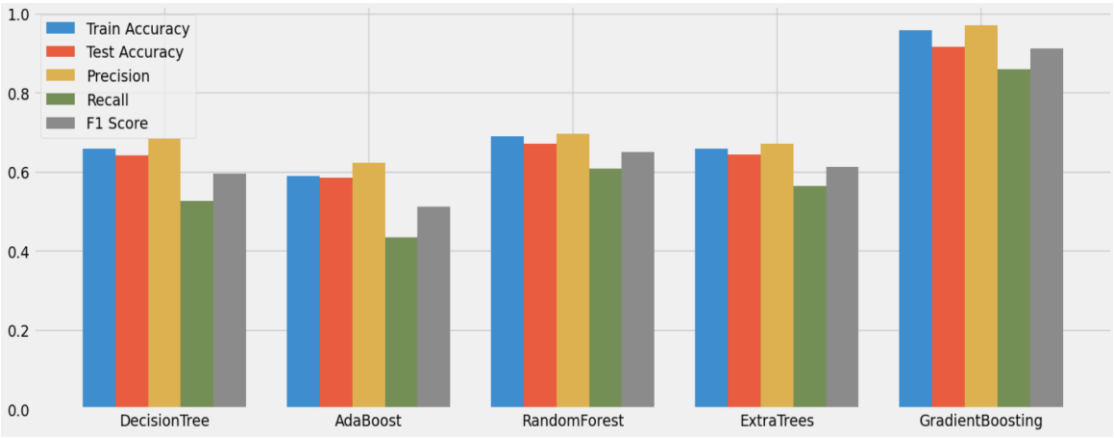

Figure 6. Graph comparing the models on different metrics

Based on the table and graph, it can be seen that the GradientBoostingClassifier model has the highest accuracy on the test dataset, as well as high scores on the other metrics, making it the preferred model for real-world applications. The results in the console, which show accuracy, F1-score, completeness, and precision equal to 1.0, are likely the result of prediction on the training dataset (trainprediction) where the model has already seen all the answers, leading to overfitting. However, in a real situation, it is important to look at the results of the test dataset (prediction), which is new data that was not involved in training the model.

3.2. Discussion of the limitations of current methods

The current study utilized computers with limited computational resources, which imposed limitations on the sample size to be processed and the complexity of the ML models used. In turn, this can lead to models trained on smaller samples not fully reflecting the complexity of the data in larger and more diverse sets, which can reduce the generalizability of the model and the accuracy of the predictions.

Limitations of the chosen model, i.e. the algorithms used, such as decision trees and random forest, although good at classification tasks, may be susceptible to overtraining and may not capture all the nuances in the data, which is particularly important in disease prediction. Suggestions for model improvement:

- Improved computational resources are justified because the use of more powerful computational hardware will allow processing of larger amounts of data, which will improve the accuracy and reliability of ML models.
- More powerful resources will also allow the use of more sophisticated models, such as deep neural networks, which can provide better performance for prediction tasks.
- Expanding the dataset. Using a larger and more diverse dataset will help improve the generalizability of models and make predictions more accurate and reliable.
- Application of dimensionality reduction techniques. Dimensionality reduction techniques such as PCA (principal component analysis) can help reduce data complexity and improve model performance on limited computational resources.

4. CONCLUSION

This paper identifies a crucial gap: the lack of integrated ML and deep learning approaches tailored for the prediction of diabetes, a burgeoning global health concern. Existing models demonstrate substantial potential, yet they require refinement to enhance predictive accuracy and to incorporate diverse patient datasets. Our study seeks to bridge this gap by unveiling non-standard correlations within diabetes prediction models, leveraging the strengths of both ML and deep learning. We aim to transcend traditional diagnostic models, probing deeper into the parameters that govern them and unveiling new diagnostic correlations.

It is important to note that the incorporation of machine and deep learning methods in the medical diagnosis of diabetes signifies a noteworthy advancement in the realm of precision medicine. Our study's findings underscore that sophisticated ML algorithms can attain elevated levels of accuracy, sensitivity, and specificity in predicting diabetes. This, in turn, enhances the quality of healthcare and augments the efficiency of clinical decision-making. Nevertheless, it is imperative to emphasize that optimal results necessitate meticulous data preparation and pre-processing, along with consideration of the unique characteristics inherent in a given clinical situation.

In conclusion, this study does not merely connect contemporary health issues with burgeoning WHO research but innovatively applies advanced computational techniques to diabetes prediction. The implications of our findings may significantly refine decision-making processes in medical practice and provide a scaffold for future research in predictive healthcare analytics.

To enhance the models and their practical application in clinical settings, it is recommended to persist in research efforts in this domain. This entails delving into extensive datasets and crafting algorithms capable of adapting to evolving conditions and patient-specific characteristics. Consequently, this project significantly contributes to the advancement of artificial intelligence methods in medicine, paving the way for new possibilities in the diagnosis and treatment of chronic diseases.

ACKNOWLEDGEMENT

This research was funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan project number AP22683316 “Application of machine learning algorithms for medical decision support systems”.

DATA AVAILABILITY STATEMENT

Access to the underlying algorithms and computational processes is obtainable by submitting a request to the corresponding author via the author's email, with the provided GitHub handle. Upon acceptance of the paper for publication, the authors are committed to establishing a dedicated repository to ensure the scientific reproducibility of the results, accessible to any interested parties. It is imperative to note that the repository itself does not house the primary dataset, as the ownership of the dataset does not belong to the authors. The data are distributed under the terms and conditions specified in the PhysioNet Credentialed Health Data Use Agreement 1.5.0. Access to the dataset can be secured via the following source: https://doi.org/10.13026/C2XW26, where the necessary procedures for obtaining the dataset can be found. Furthermore, the article explicitly references the utilization of the MIMIC III repository as a data source, which is openly available for use by any interested parties, in accordance with the repository's terms and conditions.

REFERENCES

[1] P. Xue, A. Y. Bai, Y. Jiang, and Y. L. Qiao, "WHO global strategy on digital health and its implications to China," Zhonghua yu *Fang yi xue za zhi [Chinese Journal of Preventive Medicine]*, vol. 56, no. 2, pp. 218–221, Feb. 2022, doi: 10.3760/cma.j.cn112150-20210616-00589.

[2] WHO, "Digital health," https://www.who.int/health-topics/digital-health#tab=tab_1.

[3] ITU, "ITU - the role of the global strategy on digital health 2020," https://www.itu.int/net4/wsis/forum/2023/Agenda/Session/334.

[4] WHO, "Health policy watch - digital health - big WHO ambitions but progress lags," https://healthpolicy-watch.news/digital-health-big-who-goals-slow-progress/.

[5] World Health Organization 2021, "The WHO global strategy on digital health 2020-2025," Geneva, pp. 1–60, 2021. Accessed: Oct. 27, 2023. [Online]. Available: https://www.evidentiq.com/resources/the-who-global-strategy-on-digital-health-what-you-need-to-know/

[6] WHO, "Crowell health solutions blog - WHO announces global initiative focused on digital health," https://www.crowellhealthsolutionsblog.com/2023/08/who-announces-global-initiative-focused-on-digital-health.

[7] L. Zhou, S. Pan, J. Wang, and A. V. Vasilakos, "Machine learning on big data: Opportunities and challenges," *Neurocomputing*, vol. 237, pp. 350–361, May 2017, doi: 10.1016/j.neucom.2017.01.026.

[8] T. A. Piskunova, "Primenenie intellektual'nogo analiza dannyh dlya sozdaniya sistemy reshayushchih pravil," in *XIII Vserossijskaya nauchno-prakticheskaya konferenciya «Tekhnologii Microsoft v teorii i praktike programmirovaniya*,

[9] O. G. Berestneva, "Sozdanie podsistemy prinyatiya reshenij v medicinskih informacionnyh sistemah," *Izvestiya Tomskogo Politekhnicheskogo Universiteta*, vol. 317, no. 5, pp. 194–197, 2010.

[10] E. V. Sadykova, "Information technology of medical expert decision making support systems," *Kratkie Soobshcheniya*, vol. 5, pp. 89–91, 2012.

[11] C. Naydanov, D. Palchunov, and P. Sazonova, "Development of automated methods for the critical condition risk prevention, based on the analysis of the knowledge obtained from patient medical records," in *2015 International Conference on Biomedical Engineering and Computational Technologies (SIBIRCON)*, IEEE, Oct. 2015, pp. 33–38, doi: 10.1109/SIBIRCON.2015.7361845.

[12] S. D. Gusev, N. S. Gusev, and E. N. Bochanova "Information support for the provision of high-quality medical care by using medical information systems," *Medical Information Systems*, 2016.

[13] A. B. Bagheri *et al.*, "Potential applications of artificial intelligence and machine learning on diagnosis, treatment, and outcome prediction to address health care disparities of chronic limb-threatening ischemia," *Seminars in Vascular Surgery*, vol. 36, no. 3, pp. 454–459, Sep. 2023, doi: 10.1053/j.semvascsurg.2023.06.003.

[14] E. G. Ross, N. H. Shah, R. L. Dalman, K. T. Nead, J. P. Cooke, and N. J. Leeper, "The use of machine learning for the identification of peripheral artery disease and future mortality risk," *Journal of Vascular Surgery*, vol. 64, no. 5, pp. 1515-1522.e3, Nov. 2016, doi: 10.1016/j.jvs.2016.04.026.

[15] J. Mutz and C. Lewis, "Mental health diagnosis and polygenic risk scores predict accelerated metabolomic ageing: a machine learning study," *Biological Psychiatry*, vol. 93, no. 9, p. S249, May 2023, doi: 10.1016/j.biopsych.2023.02.623.

[16] R. Gupta, S. Kumari, A. Senapati, R. K. Ambasta, and P. Kumar, "New era of artificial intelligence and machine learning-based detection, diagnosis, and therapeutics in Parkinson's disease," *Ageing Research Reviews*, vol. 90, p. 102013, Sep. 2023, doi: 10.1016/j.arr.2023.102013.

[17] H.-L. Ngan, K.-Y. Lam, Z. Li, J. Zhang, and Z. Cai, "Machine learning facilitates the application of mass spectrometry-based metabolomics to clinical analysis: A review of early diagnosis of high mortality rate cancers," *TrAC Trends in Analytical Chemistry*, vol. 168, p. 117333, Nov. 2023, doi: 10.1016/j.trac.2023.117333.

[18] M. Elhadary *et al.*, "Revolutionizing chronic lymphocytic leukemia diagnosis: A deep dive into the diverse applications of machine learning," *Blood Reviews*, p. 101134, Sep. 2023, doi: 10.1016/j.blre.2023.101134.

[19] L. Qian, J. Bai, Y. Huang, D. Q. Zeebaree, A. Saffari, and D. A. Zebari, "Breast cancer diagnosis using evolving deep convolutional neural network based on hybrid extreme learning machine technique and improved chimp optimization algorithm," *Biomed Signal Process Control*, vol. 87, p. 105492, Jan. 2024, doi: 10.1016/j.bspc.2023.105492.

[20] A. E. E. Rashed, A. M. Elmorsy, and A. E. M. Atwa, "Comparative evaluation of automated machine learning techniques for breast cancer diagnosis," *Biomed Signal Process Control*, vol. 86, p. 105016, Sep. 2023, doi: 10.1016/j.bspc.2023.105016.

[21] B. Asatryan, H. Bleijendaal, and A. A. M. Wilde, "Toward advanced diagnosis and management of inherited arrhythmia syndromes: Harnessing the capabilities of artificial intelligence and machine learning," *Heart Rhythm*, vol. 20, no. 10, pp. 1399–1407, Oct. 2023, doi: 10.1016/j.hrthm.2023.07.001.

[22] R. G. Martinez and D.-M. van Dongen, "Deep learning algorithms for the early detection of breast cancer: A comparative study with traditional machine learning," *Informatics in Medicine Unlocked*, vol. 41, p. 101317, 2023, doi: 10.1016/j.imu.2023.101317.

[23] N. Alballa and I. Al-Turaiki, "Machine learning approaches in COVID-19 diagnosis, mortality, and severity risk prediction: A review," *Informatics in Medicine Unlocked*, vol. 24, p. 100564, 2021, doi: 10.1016/j.imu.2021.100564.

[24] A. de F. Cobre *et al.*, "Diagnosis and prediction of COVID-19 severity: can biochemical tests and machine learning be used as prognostic indicators?" *Computers in Biology and Medicine*, vol. 134, p. 104531, Jul. 2021, doi: 10.1016/j.compbiomed.2021.104531.

[25] F. M. Albagmi, A. Alansari, D. S. A. Shawan, H. Y. AlNujaidi, and S. O. Olatunji, "Prediction of generalized anxiety levels during the Covid-19 pandemic: A machine learning-based modeling approach," *Informatics in Medicine Unlocked*, vol. 28, p. 100854, 2022, doi: 10.1016/j.imu.2022.100854.

[26] G. Costantini *et al.*, "Deep learning and machine learning-based voice analysis for the detection of COVID-19: A proposal and comparison of architectures," *Knowledge-Based Systems*, vol. 253, p. 109539, Oct. 2022, doi: 10.1016/j.knosys.2022.109539.

[27] J. Gao *et al.*, "MedML: Fusing medical knowledge and machine learning models for early pediatric COVID-19 hospitalization and severity prediction," *iScience*, vol. 25, no. 9, p. 104970, Sep. 2022, doi: 10.1016/j.isci.2022.104970.

[28] M. Hasan *et al.*, "Pre-hospital prediction of adverse outcomes in patients with suspected COVID-19: Development, application and comparison of machine learning and deep learning methods," *Computers in Biology and Medicine*, vol. 151, p. 106024, Dec. 2022, doi: 10.1016/j.compbiomed.2022.106024.

[29] P. M. Khaniabadi *et al.*, "Two-step machine learning to diagnose and predict involvement of lungs in COVID-19 and pneumonia using CT radiomics," *Computers in Biology and Medicine*, vol. 150, p. 106165, Nov. 2022, doi: 10.1016/j.compbiomed.2022.106165.

[30] H. Verma, S. Mandal, and A. Gupta, "Temporal deep learning architecture for prediction of COVID-19 cases in India," *Expert Systems with Applications*, vol. 195, p. 116611, Jun. 2022, doi: 10.1016/j.eswa.2022.116611.

[31] L. Xu, R. Magar, and A. Barati Farimani, "Forecasting COVID-19 new cases using deep learning methods," *Computers in Biology and Medicine*, vol. 144, p. 105342, May 2022, doi: 10.1016/j.compbiomed.2022.105342.

[32] V. V. Khanna, K. Chadaga, N. Sampathila, S. Prabhu, and R. Chadaga, "A machine learning and explainable artificial intelligence triage-prediction system for COVID-19," *Decision Analytics Journal*, vol. 7, p. 100246, 2023, doi: 10.1016/j.dajour.2023.100246.
[33] Q. Rafique *et al.*, "Reviewing methods of deep learning for diagnosing COVID-19, its variants and synergistic medicine combinations," *Computers in Biology and Medicine*, vol. 163, p. 107191, Sep. 2023, doi: 10.1016/j.compbiomed.2023.107191.
[34] S. A. Rakhshan, M. S. Nejad, M. Zaj, and F. H. Ghane, "Global analysis and prediction scenario of infectious outbreaks by recurrent dynamic model and machine learning models: A case study on COVID-19," *Computers in Biology and Medicine*, vol. 158, p. 106817, May 2023, doi: 10.1016/j.compbiomed.2023.106817.
[35] L. Zhou, C. Zhao, N. Liu, X. Yao, and Z. Cheng, "Improved LSTM-based deep learning model for COVID-19 prediction using optimized approach," *Engineering Applications of Artificial Intelligence*, vol. 122, p. 106157, Jun. 2023, doi: 10.1016/j.engappai.2023.106157.
[36] M. Ebrahimi, M. Mohammadi-Dehcheshmeh, E. Ebrahimie, and K. R. Petrovski, "Comprehensive analysis of machine learning models for prediction of sub-clinical mastitis: deep learning and gradient-boosted trees outperform other models," *Computers in Biology and Medicine*, vol. 114, p. 103456, Nov. 2019, doi: 10.1016/J.COMPBIOMED.2019.103456.
[37] T. Liu, W. Fan, and C. Wu, "A hybrid machine learning approach to cerebral stroke prediction based on imbalanced medical dataset," *Artificial Intelligence in Medicine*, vol. 101, p. 101723, Nov. 2019, doi: 10.1016/J.ARTMED.2019.101723.
[38] S. Spänig, A. Emberger-Klein, J. P. Sowa, A. Canbay, K. Menrad, and D. Heider, "The virtual doctor: An interactive clinical-decision-support system based on deep learning for non-invasive prediction of diabetes," *Artificial Intelligence in Medicine*, vol. 100, p. 101706, Sep. 2019, doi: 10.1016/J.ARTMED.2019.101706.
[39] M. M. Ghiasi and S. Zendehboudi, "Application of decision tree-based ensemble learning in the classification of breast cancer," *Computers in Biology and Medicine*, vol. 128, p. 104089, Jan. 2021, doi: 10.1016/J.COMPBIOMED.2020.104089.
[40] A. A. Kulich, "Komp'yuternaya sistema modelirovaniya kognitivnyh kart: podhody i metody," *Problemy Upravleniya*, pp. 2–16, Jun. 2010.
[41] D. Y. Strunkin and E. F. Abdrahmanov, "Patient's individual survival prediction system based on fuzzy neural," *Vrach i Informacionnye Tekhnologii*, vol. 5, 2012.
[42] A. A. Litvin and O. Y. Rebrova, "Intellektual'nyj analiz dannyh, logicheskoj regressii," *problemy zdorov'ya i ekologii*, pp. 10–17, 2016.
[43] E. Schekina, "System analysis for creation of program complexes of medical support systems (literature report)," *Journal of New Medical Technologies. eJournal*, vol. 2, pp. 1–1, Apr. 2017, doi: 10.12737/article_58f0b921cab9a8.03255229.
[44] M. A. R. Refat, Md. A. Amin, C. Kaushal, M. N. Yeasmin, and M. K. Islam, "A comparative analysis of early-stage diabetes prediction using machine learning and deep learning approach," in *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*, IEEE, Oct. 2021, pp. 654–659, doi: 10.1109/ISPCC53510.2021.9609364.
[45] K. Rajesh and V. Sangeetha, "Application of data mining methods and techniques for diabetes diagnosis," in *International Journal of Engineering and Innovative Technology (IJEIT)*, 2012.
[46] M. A. Sarwar, N. Kamal, W. Hamid, and M. A. Shah, "Prediction of diabetes using machine learning algorithms in healthcare," in *24th International Conference on Automation and Computing (ICAC) Newcastle Upon Tyne, United Kingdom*, 2018, pp. 1–6.
[47] L. Lama, O. Wilhelmsson, E. Norlander, L. Gustafsson, A. Lager, and P. Tynelius, "Machine learning for prediction of diabetes risk in middle-aged Swedish people Heliyon," *Heliyon*, vol. 7, no. 7, 2021.
[48] J. E. Bibault, P. Giraud, and A. Burgun, "Big data and machine learning in radiation oncology: State of the art and future prospects," *Cancer Letters*, vol. 382, no. 1, pp. 110–117, Nov. 2016, doi: 10.1016/J.CANLET.2016.05.033.
[49] J. Zheng, C. Xu, Z. Zhang, and X. Li, "Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, IEEE, Mar. 2017, pp. 1–6, doi: 10.1109/CISS.2017.7926112.
[50] E. E. Mokina, O. V. Marukhina, M. D. Shagarova, and I. A.Dubinina, "Using data mining method in the making of medical diagnostic decisions," *Tekhnicheskie Nauki*, vol. 5, pp. 269–274, 2016.
[51] I. Uvaliyeva, Z. Rakhmetullina, O. Baklanova, G. Györök, "The development of the staking-ensemble of methods for analyzing academic data // acta polytechnica hungarica," *Hungary*, vol. 19, no. 11, 2022, doi: 10.12700/APH.19.11.2022.11.1.
[52] D. Sharma, R. Kumar, and A. Jain, "Breast cancer prediction based on neural networks and extra tree classifier using feature ensemble learning," *Measurement: Sensors*, vol. 24, p. 100560, Dec. 2022, doi: 10.1016/j.measen.2022.100560.
[53] I. Uvaliyeva, S. Rustamov, S. Belginova, and A. Ismukhamedova, "Algorithm diagnosis of anemia on the basis of the method of the synthesis of the decisive rules," *Proceedings of the International Conference Application of Information and Communication Technologies-AICT 2019* – Baku, Azerbaijan, 2019, pp. 8-12, doi: 10.1109/AICT47866.2019.8981766.
[54] M. N. Uddin and R. K. Halder, "An ensemble method based multilayer dynamic system to predict cardiovascular disease using machine learning approach," *Informatics in Medicine Unlocked*, vol. 24, p. 100584, 2021, doi: 10.1016/j.imu.2021.100584.
[55] L. Ehwerhemuepha *et al.*, "A super learner ensemble of 14 statistical learning models for predicting COVID-19 severity among patients with cardiovascular conditions," *Intell Based Med*, vol. 5, p. 100030, 2021, doi: 10.1016/j.ibmed.2021.100030.
[56] A. Elhazmi *et al.*, "Machine learning decision tree algorithm role for predicting mortality in critically ill adult COVID-19 patients admitted to the ICU," *Journal of Infection and Public Health*, vol. 15, no. 7, pp. 826–834, Jul. 2022, doi: 10.1016/j.jiph.2022.06.008.
[57] E. E. Tripoliti, D. I. Fotiadis, and G. Manis, "Modifications of the construction and voting mechanisms of the random forests algorithm," *Data and Knowledge Engineering*, vol. 87, pp. 41–65, Sep. 2013, doi: 10.1016/j.datak.2013.07.002.
[58] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, pp. 1–14, Jun. 2017.
[59] Z. Liquan and S. Cheng, "Designing fuzzy inference system based on improved gradient descent method," *Journal of Systems Engineering and Electronics*, vol. 17, no. 4, pp. 853–857, Dec. 2006, doi: 10.1016/S1004-4132(07)60027-9.
[60] G. Yu, Y. Zhao, and Z. Wei, "A descent nonlinear conjugate gradient method for large-scale unconstrained optimization," *Applied Mathematics and Computation*, vol. 187, no. 2, pp. 636–643, Apr. 2007, doi: 10.1016/j.amc.2006.08.087.
[61] Y. Wang, Y. He, and Z. Zhu, "Study on fast speed fractional order gradient descent method and its application in neural networks," *Neurocomputing*, vol. 489, pp. 366–376, Jun. 2022, doi: 10.1016/j.neucom.2022.02.034.
[62] C. Sammut and G. I. Webb, *Encyclopedia of Machine Learning and Data Mining*, vol. 2. Boston, MA: Springer US, 2017, doi: 10.1007/978-1-4899-7687-1.
[63] B. Jan *et al.*, "Deep learning in big data analytics: A comparative study," *Computers and Electrical Engineering*, vol. 75, pp. 275–287, May 2019, doi: 10.1016/j.compeleceng.2017.12.009.
[64] F. Shan, X. He, D. J. Armaghani, and D. Sheng, "Effects of data smoothing and recurrent neural network (RNN) algorithms for real-time forecasting of tunnel boring machine (TBM) performance," *Journal of Rock Mechanics and Geotechnical Engineering*, Sep. 2023, doi: 10.1016/j.jrmge.2023.06.015.

[65] K. E. A. Kumar, D. V. Kalaga, Ch. M. S. Kumar, M. Kawaji, and T. M. Brenza, "Forecasting of COVID-19 using deep layer recurrent neural networks (RNNs) with gated recurrent units (GRUs) and long short-term memory (LSTM) cells," *Chaos Solitons Fractals*, vol. 146, p. 110861, May 2021, doi: 10.1016/j.chaos.2021.110861.
[66] H. Jiang, F. Qin, J. Cao, Y. Peng, and Y. Shao, "Recurrent neural network from adder's perspective: Carry-lookahead RNN," *Neural Networks*, vol. 144, pp. 297–306, Dec. 2021, doi: 10.1016/j.neunet.2021.08.032.
[67] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D*, vol. 404, p. 132306, Mar. 2020, doi: 10.1016/j.physd.2019.132306.
[68] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with LSTM," *Neural Computer*, vol. 12, no. 10, pp. 2451–2471, Oct. 2000, doi: 10.1162/089976600300015015.
[69] E. M. Dogo, O. J. Afolabi, N. I. Nwulu, B. Twala, and C. O. Aigbavboa, "A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks," in *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, IEEE, Dec. 2018, pp. 92–99. doi: 10.1109/CTEMS.2018.8769211.
[70] T. Schaul, I. Antonoglou, and D. Silver, "Unit tests for stochastic optimization," *arXiv preprint arXiv:1312.6055*, Dec. 2013.
[71] H. Robbins and S. Monro, "A stochastic approximation method," *The Annals of Mathematical Statistics*, vol. 22, no. 3, pp. 400–407, Sep. 1951, doi: 10.1214/aoms/1177729586.
[72] L. Bottou, "Stochastic gradient learning in neural networks," in *NeuroNîmes*, vol. 91, 1991, p. 12.
[73] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," *Journal of Machine Learning Research*, vol. 12, pp. 2121–2159, Mar. 2011.
[74] M. D. Zeiler, "ADADELTA: An adaptive learning rate method," *arXiv preprint arXiv:1212.5701*, Dec. 2012.
[75] D. M. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation." Accessed: Aug 22, 2023. [Online]. Available: https://www.researchgate.net/publication/276412348_Evaluation_From_precision_recall_and_F-measure_to_ROC_informedness_markedness_correlation
[76] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, Dec. 2014.

BIOGRAPHIES OF AUTHORS

Aigerim Ismukhamedova Additionally, she holds the position of Senior Research Associate at the D. Serikbayev East Kazakhstan Technical University (EKTU), where she is involved in project AP19679525. This project, which is funded by a grant, focuses on her work developing a "Software package for diagnostics of clinical and hematological syndromes for an electronic health passport." currently serves as the Director of the Digitalization Center at the Kazakh-American Free University (KAFU). Her educational background includes pursuing a Ph.D. in Information Systems at EKTU. Prior to that, she completed her master's degree in Computer Engineering and Software at the Almaty University of Power Engineering and Telecommunications (AUPET). Since 2017, Ismukhamedova has been conducting research as part of her doctoral studies on her Ph.D. dissertation titled "Algorithmic Support for an Intelligent Clinical Decision Support System." Under the guidance of her academic supervisor, she has published 15 scientific works related to her dissertation topic. Her work exemplifies a strong commitment to advancing the field of information systems through research and development of practical solutions to complex problems in clinical settings. She can be contacted at email: aigerim.ismukhamedova1@gmail.com.

Indira Uvaliyeva received her Engineering degree in Information Systems, followed by a Master of Engineering and Technology, and a Ph.D. in Information Systems from the D. Serikbayev East Kazakhstan Technical University (EKTU). With 14 years of professional experience, she currently holds the position of Professor and an academic title of Associate Professor. Her tenure in academia includes significant contributions to the Information Systems field, and her scientific research was notably funded by a grant from the Ministry of Science and Higher Education of the Republic of Kazakhstan under project number AP19679525. The grant supported her work on a "Software package for diagnostics of clinical and hematological syndromes for an electronic health passport." Dr. Uvalieva's scholarly work is reflected in her h-index of 8, which underscores the impact of her research in her field. Her academic journey and contributions to the scientific community embody a strong dedication to the advancement of information systems. She can be contacted at email: indirauvalieva@gmail.com.

Zhenisgul Rakhmetullina is a Lecturer in Engineering mathematics at the D. Serikbayev East Kazakhstan technical university (EKTU), Ust-Kamenogorsk, Kazakhstan. She received his B.Eng., M.Eng. and Ph.D. degrees in Al-Farabi Kazakh National University, in 1992, 2000 and 2011, respectively. She has been an Associate Professor in EKTU, Ust-Kamenogorsk, Kazakhstan since 2012. She is currently the Head of the Faculty of Basic Engineering Training. Her research interests include the field of gravitating disk, differential equations, motion of a material point, perpendicular plane, potential, and mathematical modelling. She can be contacted at email: rahmetullina@mail.ru.