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

**Aigerim Ismukhamedova, Indira Uvaliyeva**

School of Information Technology and Intelligent Systems, D. Serikbayev East Kazakhstan Technical University,

Ust-Kamenogorsk, East Kazakhstan, Kazakhstan

|  |  |  |
| --- | --- | --- |
| **Article Info** |  | **ABSTRACT** |
| ***Article history:***  Received month dd, yyyy  Revised month dd, yyyy  Accepted month dd, yyyy |  | This article centers around the development and analysis of machine learning 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 machine learning 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. |
| ***Keywords:***  Diabetes prevalence  WHO study  Machine learning  Deep learning  Modeling techniques  Healthcare resource allocation |
| *This is an open access article under the* [*CC BY-SA*](https://creativecommons.org/licenses/by-sa/4.0/) *license.* |
| ***Corresponding Author:***  Indira Uvaliyeva School of Information Technology and Intelligent Systems  D. Serikbayev East Kazakhstan Technical University 070004 Ust-Kamenogorsk, 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.

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 AI and Machine Learning (ML) has contributed to enhancing outcomes in the diagnosis, treatment, and prognosis for CLTI patients [13].

The efficacy of employing machine learning technologies in medicine is substantiated by numerous studies across various medical fields. In their examination of the diagnostic properties of machine learning algorithms in peripheral artery disease, the authors conclude that machine learning enables more precise classification and prediction of the disease [14]. Machine learning 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 machine learning methods and deep learning algorithms in diagnosing COVID-19 contributed to bringing the pandemic under control [23]–[35].

Machine learning 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 Deep Learning 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 machine learning 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 machine learning 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 machine learning 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 machine learning offer approaches to high-precision diagnosis and prognosis for breast cancer [51], [52], [53], cardiovascular disease [54], [55], and COVID-19 [56].

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. We will elucidate our methodology, combining various machine learning 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.

1. **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.

1. 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.

1. Modeling:

* A range of predictive models are delineated, including but not limited to, decision trees, random forests, and support vector machines, 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.

1. 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.

1. 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.

1. 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.

1. 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 is grounded in Diagnostic and Treatment Protocols, which establish systematic recommendations that guide medical diagnoses and treatments. These protocols align with WHO standards and offer a framework for identifying diabetes mellitus through various indicators like symptoms, glucose levels, and other clinical tests.

Our methodology involves the extraction and analysis of a comprehensive dataset from the MIMIC III database. The dataset is categorized into demographics, medical history, health parameters, drug therapy, patient records, hospitalization details, laboratory results, and treatment information. Each category is meticulously processed to ensure data integrity and relevance.

**Data Processing and Model Training.** Initial model training utilizes an expansive set of indicators, from laboratory tests to patient examination data. We deliberately allow the models to autonomously learn from the complete dataset to identify hidden correlations. Data preparation includes cleansing to remove duplications and anomalies, followed by integration to meet research objectives.

The data attributes utilized in this research are publicly listed and can be found at the MIMIC database schema overview. Data from 'chartevents', 'noteevents', 'labevents', and 'microbiologyevents' tables provide insights into patient status and treatment responses. Consolidation of patient and disease data was critical for establishing correlations necessary for effective model training.

**Optimal Model Determination.** Subsequent to data preparation, we employ a combination of machine learning 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. Machine Learning and Ensemble Methods**

Traditional Machine Learning algorithms often employ models with a linear decision boundary, as exemplified by Logistic Regression, which can be expressed as in (1):

(1)

where is the probability of the data point x being in class 1, θ is the parameter vector, and x is the feature vector. However, these models may struggle with high-dimensional data. To mitigate this, ensemble methods like Random Forests utilize multiple Decision Trees to create a more expressive model. Each tree in a Random Forest makes a prediction, and the class with the majority vote is chosen as the final prediction.

In our work, we utilized ensemble methods such as Bagging and Boosting to enhance the stability and quality of the model. Bagging, short for bootstrap aggregating, involves training multiple models on different subsets of the data and then aggregating their predictions [57]. Mathematically, if represents the prediction of the model, the aggregated prediction is in (2)

(2)

Boosting, on the other hand, sequentially trains models, each correcting the errors of its predecessor [58]. If is the prediction of the model, the combined prediction of T models is in (3):

(3)

where is the weight assigned to the model's prediction, often determined by its accuracy.

**2.3 Artificial Neural Networks (ANN)**

ANNs, particularly used in supervised learning, can be represented as a series of functions in (4):

(4)

where x is the input, y is the output, L is the number of layers in the network, is the activation function of the layer, and represents the parameters of the layer. The backpropagation algorithm is employed to train ANNs by updating these parameters to minimize the cost function, typically using a Gradient Descent algorithm. Deep learning methods, particularly Artificial Neural Networks using the Gradient Descent method, allow for the minimization of error rates [59]–[61].

The following models were selected for training and searching for the best option:

1. Convolutional Neural Network (CNN)

CNNs are adept at automatic feature extraction by applying a series of convolutional filters to the input data [62], [63]. The operation of a convolutional layer can be mathematically represented as in (5):

(5)

where \* denotes the convolution operation, is the input image, and is the filter or kernel. In practice, for discrete values, this becomes as in (6):

(6)

1. Recurrent Neural Networks (RNN)

RNNs, suitable for sequential data [64], have a defining characteristic where the output of a layer is fed back as input to the same layer [65]–[67]. The mathematical formulation for a simple RNN unit at time step is in (7) and (8):

(7)

(8)

where is the hidden state at time , is the input at time , W and b are the weights and biases, respectively, and is the output at time .

1. Long Short-Term Memory Networks (LSTM)

LSTMs enhance RNNs by incorporating memory cells that regulate the flow of information using gates [68]. The cell state Ct and hidden state at time step are updated as in (9) and (10):

(9)

(10)

where is the sigmoid activation function, , , and are the input, forget, and output gates, respectively, and is the candidate cell state.

1. Cost Function

The Cost Function, , quantifies the loss for a given set of model parameters during training [69]. It is defined as in (11):

(11)

where is the number of training examples, is the actual label, is the predicted label from the model, and is the loss measured for each prediction. In the context of ANN training, particularly for CNNs, a common choice for L is the cross-entropy loss function, defined as in (12):

(12)

where *y* is the true label and is the predicted probability. The cross-entropy function effectively captures the distance between the actual and the predicted labels, providing a mechanism to guide the gradient update during backpropagation [69].

1. Gradient Descent

Gradient Descent updates the parameters θ by moving in the direction of the negative gradient of the Cost Function. The update rule at each iteration *t* is in (13):

(13)

where η is the learning rate, and is the gradient of the cost function with respect to the parameters at iteration *t*. The learning rate η determines the step size of each update and is crucial for the convergence of the algorithm. In our work we are using a mini-batch gradient descent approach because it allows for faster convergence and less computational load per iteration [70]. The gradient is computed on a subset of the training data as is in (14):

(14)

where *m* is the mini-batch size.

The Adam optimizer further refines this approach by adjusting the learning rate for each parameter based on the estimates of the first and second moments of the gradients. Mathematical representation is in (15):

(15)

where and are the bias-corrected first and second moment estimates, respectively, and ϵ is a small scalar used to prevent division by zero [71]–[74].

**2.4 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 (16):

(16)

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.

**2.5. Evaluation Metrics for Classification Models**

In the quantitative analysis of classification models within the realms of Machine Learning and Deep Learning, several performance evaluation metrics stand as pillars for assessing algorithmic efficacy. Key among these are precision, recall [75], F1-score [75], [76], accuracy, and the confusion matrix.

The confusion matrix, illustrated in Figure 1, offers a visual and numerical representation of the predictive capacity of a classification algorithm. It delineates the frequency of each class prediction against actual labels, providing a foundational framework for further metric calculations.

The matrix is partitioned into four quadrants reflecting the counts of:

* True Positives (TP): Instances correctly predicted as positive,
* True Negatives (TN): Instances correctly predicted as negative,
* False Positives (FP): Instances incorrectly predicted as positive, often referred to as Type I error,
* False Negatives (FN): Instances incorrectly predicted as negative, often referred to as Type II error.

A table with text on it

Description automatically generated

Figure 1. Confusion Matrix or Contingency table

These quadrants enable us to compute the following metrics:

1. Accuracy

Accuracy quantifies the overall correctness of the model and is calculated as the sum of true predictions over all predictions. It is presented in (17):

(17)

1. Precision

Precision, also known as positive predictive value, measures the proportion of true positives among all positive predictions. It is presented in (18):

(18)

1. Recall (Sensitivity)

Recall, or sensitivity, assesses the proportion of actual positives that are correctly identified. It is presented in (19):

(19)

1. F1-score

The F1-score harmonizes precision and recall, providing a single measure of a test’s accuracy. It is presented in (20):

(20)

To systematize these computations, one may define a vector y representing actual class labels and a vector 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 (21):

(21)

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.

**2.6. Data preparation and pre-processing**

Data were extracted from the MIMIC-III database, comprising medical records of numerous patients. The study focused on records related to diabetic patients, encompassing demographic data, medical history, medication records, and hospitalizations. Consequently, the training data included information on the entire medical history and tests, not exclusively limited to diabetes-related records. The absence of specificity in the training data, in our opinion, is a crucial factor for uncovering additional hidden dependencies and identifying new predictive parameters.

Data cleansing entails the removal of incomplete or inaccurate records. For instance, if a patient record lacked essential information regarding diagnosis or treatment, it was excluded from the analysis. Managing missing values becomes crucial when data are absent; for instance, if information such as a patient's age or gender is missing, imputation techniques were employed. Alternatively, records with missing values were excluded from the analysis to uphold data accuracy.

We employed the standardization of numerical data, utilizing techniques like the Z-transform to homogenize the scales of numerical data. For instance, laboratory test scores underwent standardization to guarantee the comparability of results.

Another critical aspect was the transformation of categorical data. Encoding categorical variables like gender or race was carried out using one-hot encoding methods to ensure proper processing by machine learning models.

Data preprocessing is a crucial step aimed at eliminating biases and errors that may adversely affect the accuracy and reliability of machine learning models. Through data standardization and transformation, we ensured the homogeneity and comparability of data, which are essential for the analysis and interpretation of results. These steps guaranteed high data quality, a pivotal factor in obtaining accurate and meaningful study results.

The model selection process was driven by the characteristics of the data, including factors such as sample size, the number of features, and the type of task. Models well-suited for the diabetes classification task were chosen.

Decision tree, random forest, and AdaBoost models were employed. Decision trees were preferred for their interpretability, random forests for their robustness to overfitting, and AdaBoost for enhancing the performance of weak classifiers. Random forest was specifically selected due to its efficient handling of large datasets and numerous features, as well as its robustness to overfitting [59].

The hyperparameter tuning process employed methods such as Grid Search or Random Search to identify the optimal hyperparameters for the models.

The selected hyperparameters included the following:

* For the random forest, the number of trees, tree depth, and splitting criteria were considered.
* In AdaBoost, the crucial hyperparameters were the number of iterations and the learning rate.

Tuning hyperparameters is crucial for enhancing model performance and mitigating both overfitting and underfitting. Utilize the inherent capabilities of machine learning models to gauge the importance of features, such as the feature\_importances\_ attribute in a random forest. For instance, identifying the most crucial features for predicting diabetes, such as blood glucose level, BMI, age, and gender, can be accomplished. Understanding the significance of features aids in model interpretation, enabling a focus on the most substantial aspects of the data.

1. **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:

1. imported all the necessary libraries:
   1. pandas
   2. numpy
   3. matplotlip
   4. seaborn
   5. sklearn
2. 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:
   1. Number of Patients with Diabetes: (552336, 13)
   2. Number of Patients without Diabetes: (15969, 13)
   3. Number of Final balanced Dataframe: (31938, 13)
3. Next, we begin our modeling
   1. Importing the necessary libraries. The code starts by importing classes from the sklearn library used for cross-validation and various classification algorithms.
   2. 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.
   3. Creating a list of classifiers. Various classification algorithms are added to the list of classifiers, including DecisionTreeClassifier and ensemble methods such as AdaBoostClassifier, RandomForestClassifier, ExtraTreesClassifier, and GradientBoostingClassifier.
   4. 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.
   5. For each classifier, the average accuracy and standard deviation of the cross-validation results are computed.
   6. 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.

**Изображение выглядит как снимок экрана, текст, линия, График

Автоматически созданное описание**

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

The diagram shows that GradientBoostingClassifier has the highest average accuracy while DecisionTreeClassifier has the lowest. This suggests that ensemble learning based models generally outperform the simple DecisionTreeClassifier, 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:

1. 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.
2. 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.

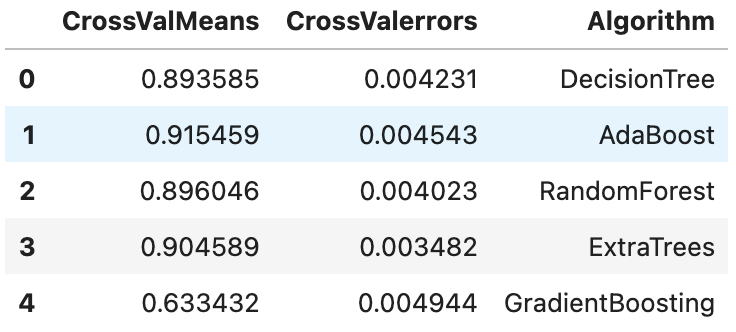
****

Figure 3. Comparative Performance of Classification Algorithms

From the data provided, we can infer the following:

1. 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.
2. 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.
3. 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.

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

a. Initialization of EstimatorSelectionHelper (EstimatorSelectionHelper). The constructor accepts dictionaries with models and their corresponding hyperparameters. It is checked that parameters are provided for each model.

b. 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.

c. 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.

d. 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.

e. 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 GradientBoostingClassifier 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 AdaBoostClassifier and DecisionTreeClassifier in this setting.

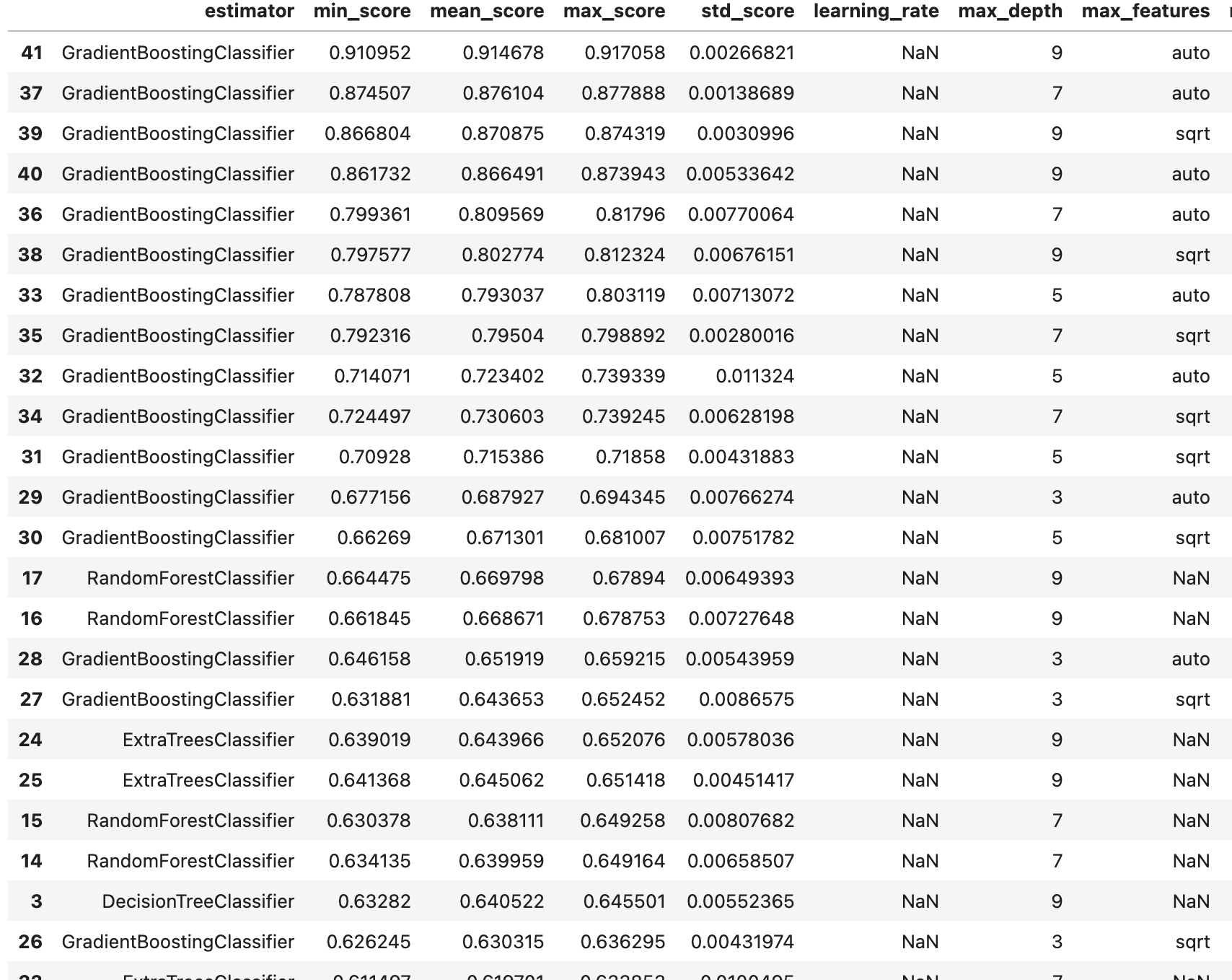
****

Figure 4. Results of models and selected hyperparameters

5) 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:

1. Importation of the metrics library. The code commences with the importation of functions for computing the model's quality metrics.
2. 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.
3. 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).
4. Calculation of metrics. For each model, accuracy, precision, recall, and F1-score are computed both on the training and the test datasets.
5. Reporting of results. The calculated metrics for each model are outputted, along with a classification report and confusion matrix.

The calculated metrics, confusion matrix (22) and classification report (Table 1) for DecisionTree model are represented below:

* Accuracy: 1.0
* F1 score: 1.0
* Recall: 1.0
* Precision: 1.0

(22)

Table 1. The classification report for DecisionTree 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 (23) and classification report (Table 2) for AdaBoost model are represented below:

* Accuracy: 1.0
* F1 score: 1.0
* Recall: 0.0
* Precision: 1.0

(23)

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 (24) and classification report (Table 3) for RandomForest model are represented below:

* Accuracy: 1.0
* F1 score: 1.0
* Recall: 1.0
* Precision: 1.0

(24)

Table 3. The classification report for RandomForest 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 (25) and classification report (Table 4) for ExtraTrees model are represented below:

* Accuracy: 1.0
* F1 score: 1.0
* Recall: 1.0
* Precision: 1.0

(25)

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 (26) and classification report (Table 5) for GradientBoosting model are represented below:

* Accuracy: 1.0
* F1 score: 1.0
* Recall: 1.0
* Precision: 1.0

(25)

Table 5. The classification report for GradientBoosting 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 |

* 1. Creating and outputting a summary table, represented in Figure 5. It collects information on each metric for all models.

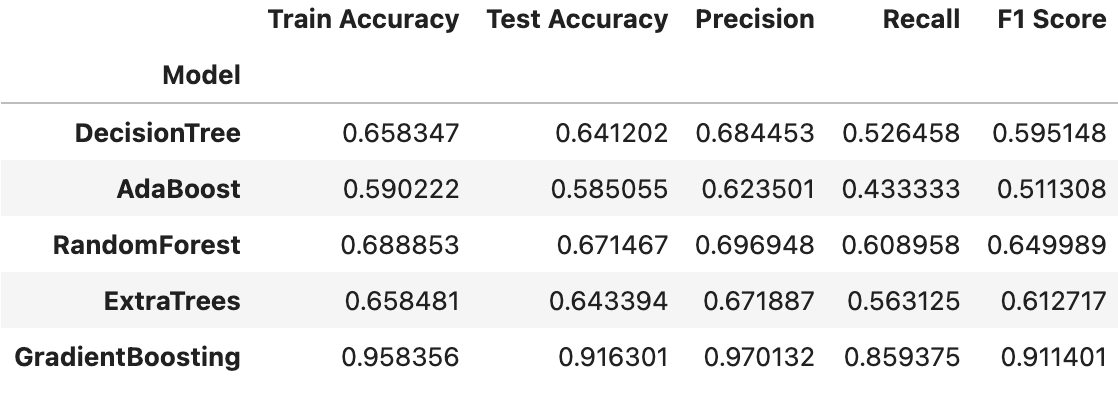
****

Figure 5. Summary table of each metric for all models

* 1. 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.

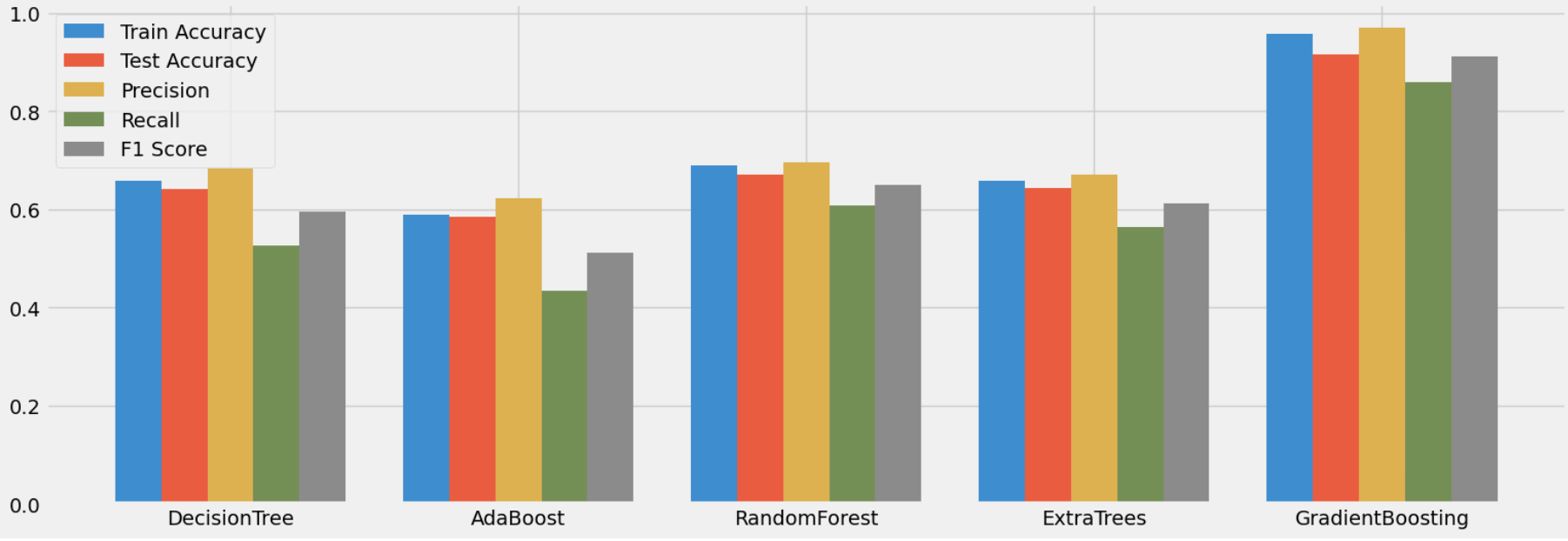
****

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 machine learning 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 machine learning 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.

1. **CONCLUSION**

This paper identifies a crucial gap: the lack of integrated machine learning 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 machine learning 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 machine learning 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.

**FUNDING**

This research was funded by Ministry of Science and Higher Education of the Republic of Kazakhstan project number AP19679525 “software package for diagnostics of clinical and hematological syndromes for electronic health passport”.

**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,” 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] Piskunova T.A, “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] Gusev S.D, Gusev N.S, and Bochanova E.N., “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,” *Semin Vasc Surg*, 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,” *J Vasc Surg*, 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,” *Biol 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 Res Rev*, 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 Rev*, 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. Eldin Rashed, A. M. Elmorsy, and A. E. Mansour 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. Gonzales Martinez and D.-M. van Dongen, “Deep learning algorithms for the early detection of breast cancer: A comparative study with traditional machine learning,” *Inform Med 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,” *Inform Med 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?,” *Comput Biol Med*, vol. 134, p. 104531, Jul. 2021, doi: 10.1016/j.compbiomed.2021.104531.

[25] F. M. Albagmi, A. Alansari, D. S. Al Shawan, H. Y. AlNujaidi, and S. O. Olatunji, “Prediction of generalized anxiety levels during the Covid-19 pandemic: A machine learning-based modeling approach,” *Inform Med 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,” *Knowl Based Syst*, 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,” *Comput Biol Med*, vol. 151, p. 106024, Dec. 2022, doi: 10.1016/j.compbiomed.2022.106024.

[29] P. Moradi Khaniabadi *et al.*, “Two-step machine learning to diagnose and predict involvement of lungs in COVID-19 and pneumonia using CT radiomics,” *Comput Biol Med*, 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 Syst Appl*, 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,” *Comput Biol Med*, 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. C. P., “A machine learning and explainable artificial intelligence triage-prediction system for COVID-19,” *Decision Analytics Journal*, vol. 7, p. 100246, Jun. 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,” *Comput Biol Med*, 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,” *Comput Biol Med*, 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,” *Eng Appl Artif Intell*, 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,” *Comput Biol Med*, 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,” *Artif Intell Med*, 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,” *Artif Intell Med*, 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,” *Comput Biol Med*, vol. 128, p. 104089, Jan. 2021, doi: 10.1016/J.COMPBIOMED.2020.104089.

[40] Kulich A. A., “Komp’yuternaya sistema modelirovaniya kognitivnyh kart: podhody i metody,” *obzory. обзоры* , pp. 2–16, Jun. 2010.

[41] Strunkin D. YU and Abdrahmanov E. F., “Patient’s individual survival prediction system based on fuzzy neural,” *Vrach i informacionnye tekhnologii*, vol. 5, 2012.

[42] A.A. Litvin and O.YU. 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. Al 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] Rajesh K. and Sangeetha V., “Application of data mining methods and techniques for diabetes diagnosis,” in *Int. J. Eng. Innov. Technol. (IJEIT)*, 2012.

[46] Sarwar M.A., Kamal N., Hamid W., and Shah M.A., “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] Lama L., Wilhelmsson O., Norlander E., Gustafsson L., Lager A., and Tynelius P., “Machine learning for prediction of diabetes risk in middle-aged Swedish people 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 Lett*, vol. 382, no. 1, pp. 110–117, Nov. 2016, doi: 10.1016/J.CANLET.2016.05.033.

[49] Jian Zheng, Cencen Xu, Ziang Zhang, and Xiaohua 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] Mokina Е.Е, Marukhina О.V, Shagarova M.D, and Dubinina I.A., “Using data mining method in the making of medical diagnostic decisions,” *tekhnicheskie nauki*, vol. 5, pp. 269–274, 2016.

[51] Indira Uvaliyeva, Zhenisgul Rakhmetullina, Olga Baklanovа, György 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] Uvaliyeva I., Rustamov S., Belginova S., Ismukhamedova A. 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,” *Inform Med 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,” *J Infect 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 Knowl Eng*, 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,” 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,” *Appl Math Comput*, 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 & 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. ArunKumar, 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 Comput*, 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,” 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,” 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,” Dec. 2014.

**BIOGRAPHIES OF AUTHORS (10 PT)**

|  |  |
| --- | --- |
|  | **Ismukhamedova Aigerim**   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 |
|  |  |
|  | **Uvalieva Indira Makhmutovna** 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 |
|  |  |