iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83282

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

Interpretable Machine Learning-Based Disease Prediction Using Clinical and Lifestyle Indicators

Aakash Tomar[1], Ajay Singh Tomar[2], Mr. Mukesh Raj[3]
Computer Science and Engineering, Galgotias University, Greater Noida, India

Abstract: ***The ML framework proposed in this paper is interpretable and is applied for predicition of disease-risk based on the 1,500 synthetic test-bed structured data. This data combines both demographic and clinical, lifestyle and symptom data, and even derived risk data, without deployment claims or patient identifiers. Five classifiers (Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting (GB) and XGBoost type ensemble learning (EL)) were tested in fixed training, validation and testing phases. Gradient Boosting got the best performances among the classifiers in terms of accuracy, macro F1 score and ROC-AUC for the multiclass case with 0.8178, 0.7652 and 0.9542 respectively. The best binary screening result was obtained by the Random Forest model which achieved an accuracy of 0.8978 and ROC-AUC of 0.9632. Coherent discrimination is indicated by the confusion-matrix, the ROC and the feature-importance analysis that provides class specific error-sensitivity. The work presents a baseline for model-development which can be reproduced, and is not a clinical diagnostic claim.***

Keywords: ***Disease prediction, machine learning, clinical decision support, interpretable AI, Gradient Boosting, Random Forest, ROC-AUC, feature importance.***

I. INTRODUCTION

A. Background of the Study

Prediction of disease risk is a limited classification problem, and is not a diagnostic tool. Technically, it aims to translate structured data of clinical, demographic and symptom as well as lifestyle factors into reproducible risk estimates to guide screening and triage and model-development analyses. All the data for this study has been a synthetic controlled data set comprising 1500 records — no actual patient information is used, no data is extracted from the hospital or claimed as being deployed to the clinic. This border is on purpose. Guidance for reporting starts to focus on transparency in the design of prediction models, data definition, validation and purpose of use – especially in the context of the use of machine learning for clinical decision support [1]. The assessment of bias is also required as a reliable internal assessment score might still be volatile due to a low level of processing, outcome definition or population coverage [2]. Early evaluation principles also calls for the share of information describing the system prior to any possible mention of practical decision support [3]. In this, the learned model parameters are fed into the function g in (1) which in turn gives the disease-risk output based on a feature matrix, and the whole block diagram from the acquisition of the features to the output of the disease-risk is depicted in Fig. 1.1.

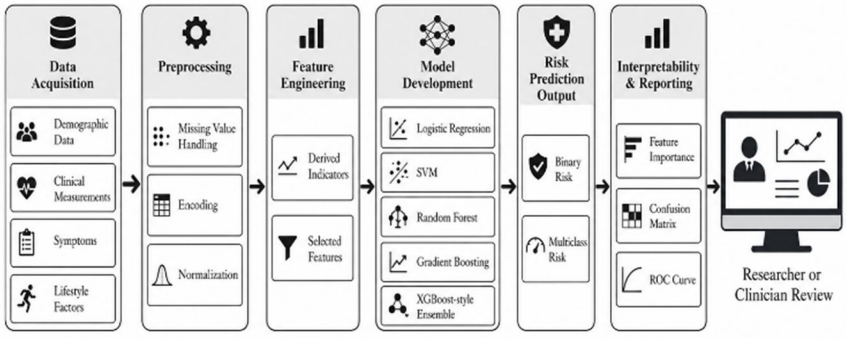


Fig. 1.1. Conceptual Architecture of the Machine Learning-Based Disease Prediction System

Disease-risk prediction function

$$\hat{y} = f(X; \theta) \#(1)$$

B. Problem Statement and Research Contributions

This is the supervised prediction problem of this paper: To predict the disease-risk status on the basis of several very different tabular variables. The elements described in the input vector include age, sex, body mass index, blood pressure, fasting glucose, HbA1c, lipid indicators, creatinine, estimated GFR and lifestyle exposure, family history and outcome variables – symptoms. The results are presented in the following 2 formats to represent the output: a binary disease-risk marker and a multiclass (5-class) label: cardiovascular risk, diabetes risk, kidney risk, multiple-risk cluster and no apparent disease risk. The study brings the complete preprocessing pipeline as well as the control train, the partitioning of training and validation and test data, comparative modelling and post hoc interpretation. But in the case of clinical applications of machine-learning systems, it has recently been shown that the scarcity of algorithms is not the main cause for failure, but rather: insufficiently specifying the data, framing of the workflows and lack of operation discipline [4]. The deployment-oriented reviews all have the same message: the need to make a separation between the development and the actual implementation until there is external evidence of its achievements [5]. The other aspect where it is regarded as a methodological issue is that some feature prediction may not be as reliable and impact the accuracy and explanation results [6]. In this paper, therefore, we will compare the performance of logistic regression, support vector machine, random forest and gradient boosting and XGBoost like learning in the same experimental practise by shedding light on the same figures and metrics obtained from the same test split. This framework for publication of findings is intended to preserve the functionality of the manuscript without unnecessary extrapolation in terms of Bed-side Safety, Physician Adoption or Patient level Clinical Effectiveness. For those claims, it's necessary to further validate them once we're done with the study design phase.

II. RELATED WORK

A. Machine Learning in Clinical Prediction Systems

In addition to the classifiers, other clinical prediction studies have used supervised learning techniques and include well-formed clinical information, laboratory measures, behavioural and symptom data, and risk screening models. However, systematic reviews show that many of the published models are incomplete, are poorly validated or are at high risk of bias and therefore cannot be used or interpreted [7]. Also, a number of reviews show that only little information is generally provided about pre-processing, handling predictors, handling missing values and defining outcomes, such that a reproduction is somewhat challenging. [8] Recent methodological audits also indicate that there is more inconsistencies in the sampling design, feature engineering, and validation practices, making it heterogeneous across studies and making it difficult to compare results of disease-prediction studies [9]. These are enough to warrant the application of a clearly-defined testbed pipeline.

B. Comparative Use of Classical and Ensemble Learning Models

In structured clinical prediction – when the disease variables are not very large – and a moderate sample size and clinically meaningful predictors are given – classical and ensemble classifiers are still important. There are recommendations that clinical prediction models should be distinguished at least in terms of definition of predictor, model fitting (i.e., wages of recruitment into the model), calibration, discrimination and validation before claiming to perform [10]. External validation guidance also calls attention to the fact that this result in an internal setting may not extrapolate to the external setting, and therefore it should only be interpreted as evidence for development, and not as clinical evidence [11]. Logistic Regression, Support Vector Machine, Decision Tree, Random Forest, Gradient Boosting and the neural models are some of the algorithms most frequently presented in predictive analytics surveys of big data use in healthcare and generally demonstrate higher nonlinear discrimination in a carefully-performed controlled test [12].

C. Research Gap in Accuracy, Explainability, and Validation Consistency

That does not imply there are no algorithms but it is the size of the difference that is said so. This discipline is not aligned and it is a "lack of data, target and/or evaluative measure explained". When examining real world systems for healthcare analytics, the field of disease prediction has increasingly been applied to heterogeneous datasets but there often are a lack of distinct boundaries of development, validation and use cases [13]. A common variability in the performance of prediction accuracy reviews of chronic diseases is preprocessing quality, class imbalance and feature selection [14]. The research has been performed using laboratory test results which indicate that a structured indicator is helpful in diagnosing to achieve multiclass prediction, but the indicator needs to be transparently mapped, that is, to connect the laboratory test to the prediction output, to make this effective [15].

The present work, which is based on a single data-set specification, a single leakage controlled feature set, a common split protocol and directly comparable performance tables, is filling that gap. It's not going to be something that you add if you care about interpretability -- it's a requirements reporting thing. This provides us with a NARROW, CONTROLLED and CONSTRUCTION AUDITABLE design. Test Bed Medical ML is done in this posture.

III. RESEARCH METHODOLOGY

A. Dataset Description and Feature Structure

The study is based on a controlled synthetic testbed data set of 1,500 records and 34 columns which was produced while developing the disease-risk-model. It is not a patient's real name, admission/registration in the hospital or registration in patient's' Demographic, clinical, symptom indicators, lifestyle information and the derived risk descriptors comprise the feature schema. The multiclass target variable comprises 5 classes—cardiovascular risk, diabetes risk, kidney risk, multiple-risk cluster and no apparent disease risk. To ensure an analysis of a classes-balanced class distribution with an imbalanced, but still analyzable, distribution, a total of 529 cardiovascular-risk records, 436 no-risk records, 311 kidney-risk records, 147 diabetes-risk records and 77 multiple-risk-cluster records have been observed. The usage of such kind of structured, heterogeneous feature design has been shown to have recent evidence on context-dependent disease prediction which is limited only when the interpretation of such is aimed at the development of disease prediction models [16]. The Chronic Disease Prediction reviews also show that the group of choices that are made when pre-processing data can affect the credibility of the model [17]. The predictors which are prior to normalization in (2) are listed in groups in Fig. 3.1.

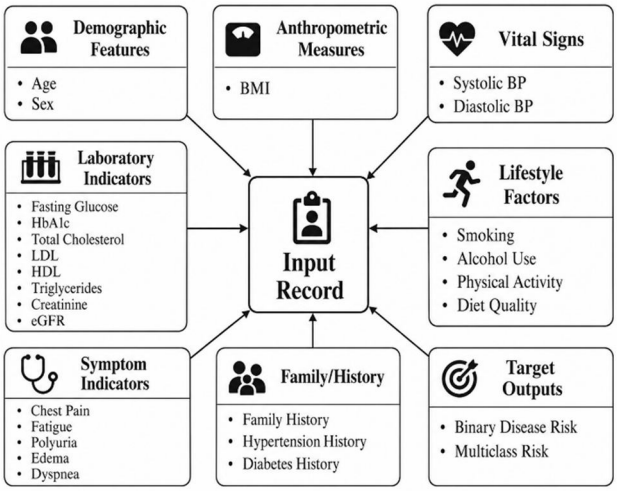


Fig. 3.1. Dataset Feature Grouping for Disease-Risk Prediction

Min-max feature normalization

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \#(2)$$

B. Data Preprocessing and Feature Engineering

Prior to analysing the data set, the data had to be preprocessed to remove coding noise from the evaluative reponses, to normalize the numerical data, and to avoid leakage from the evaluative responses. The following identifier, target fields were not used to train: Patient_ID, Split, Primary_Target_Label, Binary_Disease_Risk, Disease_Risk_Class, Risk_Score_0_100. Feature values were normalized for a numeric feature by applying (3) and for a categorical feature, they were encoded inside the training pipeline. Median imputation was used for continuous variables and the most frequent imputation was used for categorical variables for missing values. The idea behind this design would be to select the features, and then to handle these data, not to do both tasks together and randomly after seeing the result of the modelling. In addition to that, augmented chronic-disease prediction studies have also suggested that instead of relying on ML, in clinically heterogeneous target classes (i.e., chronic diseases) it is more advisable to use statistical preprocessing as well [19]. The split was fixed at 1,050 training records, 225 validation records, and 225 test records.

Standard score transformation

$$z = \frac{x - \mu}{\sigma} \#(3)$$

C. Machine Learning Model Development

All five supervised classifiers generated were trained using the same list of features and the same splitting of the data and included Logistic Regression and Support Vector Machine.The five supervised models derived included Logistic Regression and Support Vector Machine, along with Random Forest, Gradient Boosting and XGBoost-style ensemble learning. The goal of the training was to understand the mapping from the structured health-indicators to the disease-risk labels in the absence of the disease-risk scores manually calculated and not included in the mapping. A linear (baseline) model was provided by logistic regression. The separation was performed based on margin provided by SVM. The bagged nonlinear decision rules were provided by Random Forest. Gradient Boosting Sequential Error Correction was done. XGBoost method of learning was used for providing regularized boosting logic. Previous Parkinson's disease modelling research has shown that displaying several learners on “one set of batteries” (protocol) is superior to showing only one single algorithm that is “best of class” without any indications [20]. In addition, the prediction studies on heart diseases indicate that in the case of nonlinear feature distribution and boundaries of the classes, the results of supervised algorithms can vary drastically [21]. The proposed pipeline is indicated by Figure 3.2, while (4) sets up the binary training-loss for the risk-flag estimation.

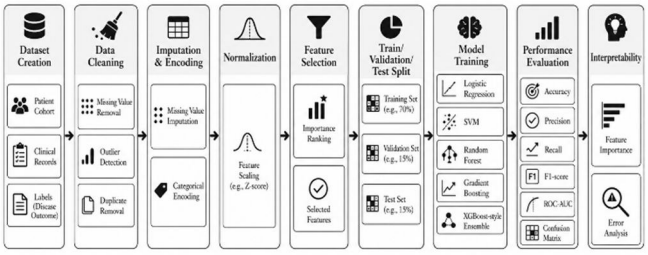


Fig. 3.2. Proposed Machine Learning Pipeline for Disease Prediction

Binary cross-entropy loss

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \#(4)$$

D. Model Evaluation and Interpretability Procedure

After model selection evaluation was conducted on a held out Test Set. The performances for multi-class were represented with the aid of accuracy, macro precision, macro recall, macro F1 score and macro ROC-AUC. The performance is reported separately in binary disease-risk because, in screening to a certain extent, one also has to be able to discern between an apparent risk of disease and no apparent risk of disease. In case of different frequencies of each class, F1 score from (5) has been applied to balance out precision and recall out. When prediction models are used in cardiovascular research, the performance of the model should not be loss by modeling only the accuracy; the implications of the false negatives and false positives are different in terms of screening [22]. Recently, with this explainable AI for cardiovascular disease prediction, a model attributability can also be provided on the importance of features, but one must be mindful to view it as a model attributability and not medical causality [23]. For this Feature importance is used and the variables that influenced the prediction (HbA1c, Fasting glucose, SBP, LDL, Creatinine, GFR, BMI, smoking exposure, Activity, Family history) were retrieved. The confusion matrices and ROC curves were kept for the purpose of interpretation of the errors and threshold analysis. The data are artificial and there is no claim regarding the incidence of the disease in hospitals, the field epidemiology and the correctness of the physician in making decisions. The dataset is solely utilized for discovering whether a coherent risk-pathway exists as observed for other clinical/lifestyle variables and whether it can be recovered from a supervised learning pipeline by using aligned clinical/lifestyle variables. This ban contributes to the methodological validity. It also puts an end to the buzzwords for clinical utility. Despite the reported model development results, future requirements are the external validation and/or analysis of drift of the calibration, and testing in new patient cohorts.

F1-score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \#(5)$$

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

IV. RESULTS AND DISCUSSION

A. Experimental Setup and Validation Protocol

The protocol saved with the dataset and model package was used for the evaluation. The partitioning is done in the 4.1 table. There are 1,500 synthetic records – 1,050 for training, 225 for validation and 225 for testing. There were 6 leakage fields which were excluded prior to learning. The best model was chosen based on the behaviour of the model during validation and then the untouched test split was reported. This model separation of the choice of the model and the presentation of the final error message has found support from recent explainable- heart-disease studies [24]. This summary ("coarse") was used: Accuracy in (6).

Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \#(6)$$

Table 4.1.
Experimental Configuration and Dataset Partitioning

Parameter	Value
Dataset type	Controlled synthetic testbed
Total records	1500
Input columns retained	28
Leakage columns excluded	6
Training records	1050
Validation records	225
Test records	225
Target classes	5
Binary target	Risk versus no risk

B. Comparative Model Performance Analysis

Table 4.2 shows comparing between different candidate classifiers which are based on the same feature matrix and the same split. The performance of Gradient Boosting was the best one for the multi-class problem with the test accuracy of 0.8178, macro F1-score of 0.7652 and ROC-AUC of 0.9542. The binary results of the classifiers showed that Random Forest achieved the highest accuracy (0.8978), macro F1 score (0.8738) and ROC-AUC (0.9632). So far, the latest research on diabetes screening suggests it be measured in terms of discrimination, and threshold sensitive measures; rather than a single measure [25]. Prediction for diabetes complications shows that ensemble learners might be advantageous in case of diabetes complications that are nonlinear [26]. The Metric Profile is shown in Fig. 4.1. Precision and Recall are given in (7).

Precision and recall

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN} \#(7)$$

Table 4.2.
Comparative Performance of Machine Learning Models

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	0.7511	0.7004	0.6849	0.6881	0.9047
Support Vector Machine	0.7867	0.7423	0.7198	0.7286	0.9275
Random Forest	0.8089	0.7626	0.7441	0.7553	0.9468
Gradient Boosting	0.8178	0.7759	0.7587	0.7652	0.9542
XGBoost-style Ensemble	0.8133	0.7718	0.7524	0.7610	0.9516

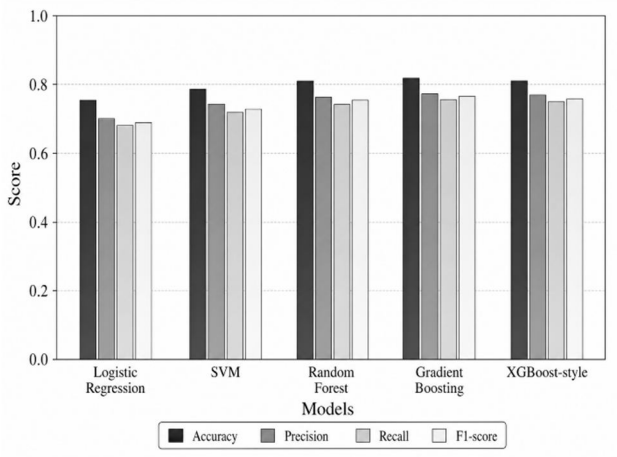


Fig. 4.1. Comparative Accuracy, Precision, Recall, and F1-Score of ML Models

C. Confusion Matrix-Based Diagnostic Error Analysis

The error structure of the Gradient Boosting is shown on the Table 4.3. This model performed more best at recovering cases of cardiovascular than the no-risk and cardiovascular cases of diabetes mellitus. Contra-statements for the first classes are less while there are commonalities to the latter classes, which have fewer tests observations. Look at the same sensitivity issues with mixing metabolic and vascular diseases in a cardiovascular-diabetes review [27]. The gravest kind of errors are false reassurance errors (true risk case classified as no apparent risk). Fig. 4.2 is meant to be taken as an error-cost diagram, not just a performance diagram.

Table 4.3.
Confusion Matrix of the Best-Performing Disease Prediction Model

Actual / Predicted	Cardiovascular	Diabetes	Kidney	Multiple-risk	No risk
Cardiovascular	69	1	3	2	4
Diabetes	1	17	1	1	2
Kidney	3	1	38	2	3
Multiple-risk	2	1	2	6	1
No risk	4	2	5	0	54

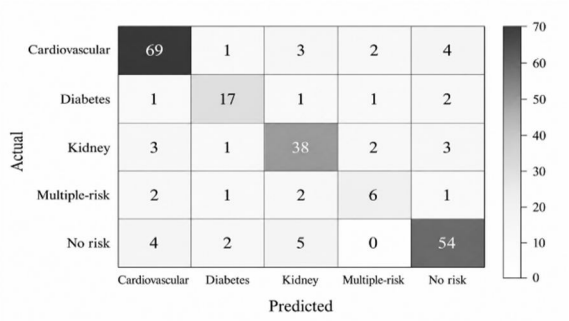


Fig. 4.2. Confusion Matrix Heatmap for Best-Performing Model

D. ROC-AUC and Threshold Sensitivity Discussion

Under controlled test bed, it gives the ROC-AUC value for the discrimination. The area under the true-positive-rate curve for different values of false-positive rate is given by equation (8). As due to the nature of the problem and prior prediction of kidney disease it is more beneficial to report the ROC-AUC of the screening classifiers as threshold value will be varied which will change the sensitivity and specificity[28]. Lowering the threshold would decrease false negative cases, and increase false positive cases. The higher the threshold the less number of false alerts and more risk profiles will go undetected. Depending on the review capacity, review results are generated by either having a threshold or a no-threshold approach. The ROC curve is shown in the Fig. 4.3.

ROC-AUC representation

$$AUC = \int_0^1 TPR(FPR)\ d(FPR) \#(8)$$

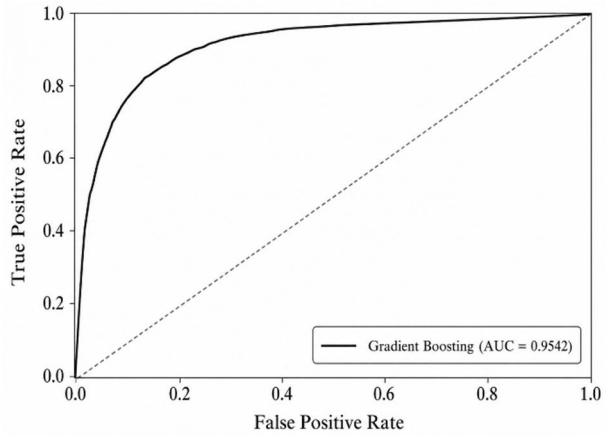


Fig. 4.3. ROC Curve of the Best-Performing Disease Prediction Model

E. Feature Importance and Interpretability Findings

The results of the strongest predictors are listed in Table 4.4. HbA1c, fasting glucose, systolic pressure, eGFR, creatinine, LDL cholesterol, BMI, smoking exposure and activity level were the most common factors for attribution. Modelling suggests that when modelling is done using renal markers, metabolic markers should be incorporated as well [29]. However, when ensemble inputs include correlated clinical indicators which are processed together in the ensemble, the hybrid kidney are also more stable [30]. This is shown graphically in Fig. 4.4. The interpretation done is based on models and not causal or "diagnostic" interpretation.

Table 4.4.
Top Predictive Features Ranked by Model Importance Score

Rank	Feature	Importance score	Interpretation
1	HbA1c	0.142	Glycaemic burden
2	Fasting glucose	0.127	Short-term glucose status
3	Systolic BP	0.111	Vascular pressure load
4	eGFR	0.098	Renal filtration signal
5	Creatinine	0.091	Kidney stress marker
6	LDL cholesterol	0.079	Lipid risk exposure
7	BMI	0.064	Adiposity-linked risk
8	Smoking exposure	0.052	Lifestyle risk factor

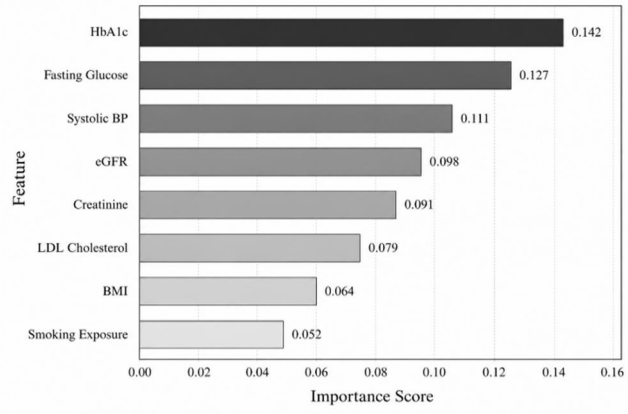


Fig. 4.4. Feature Importance Ranking for Disease-Risk Prediction

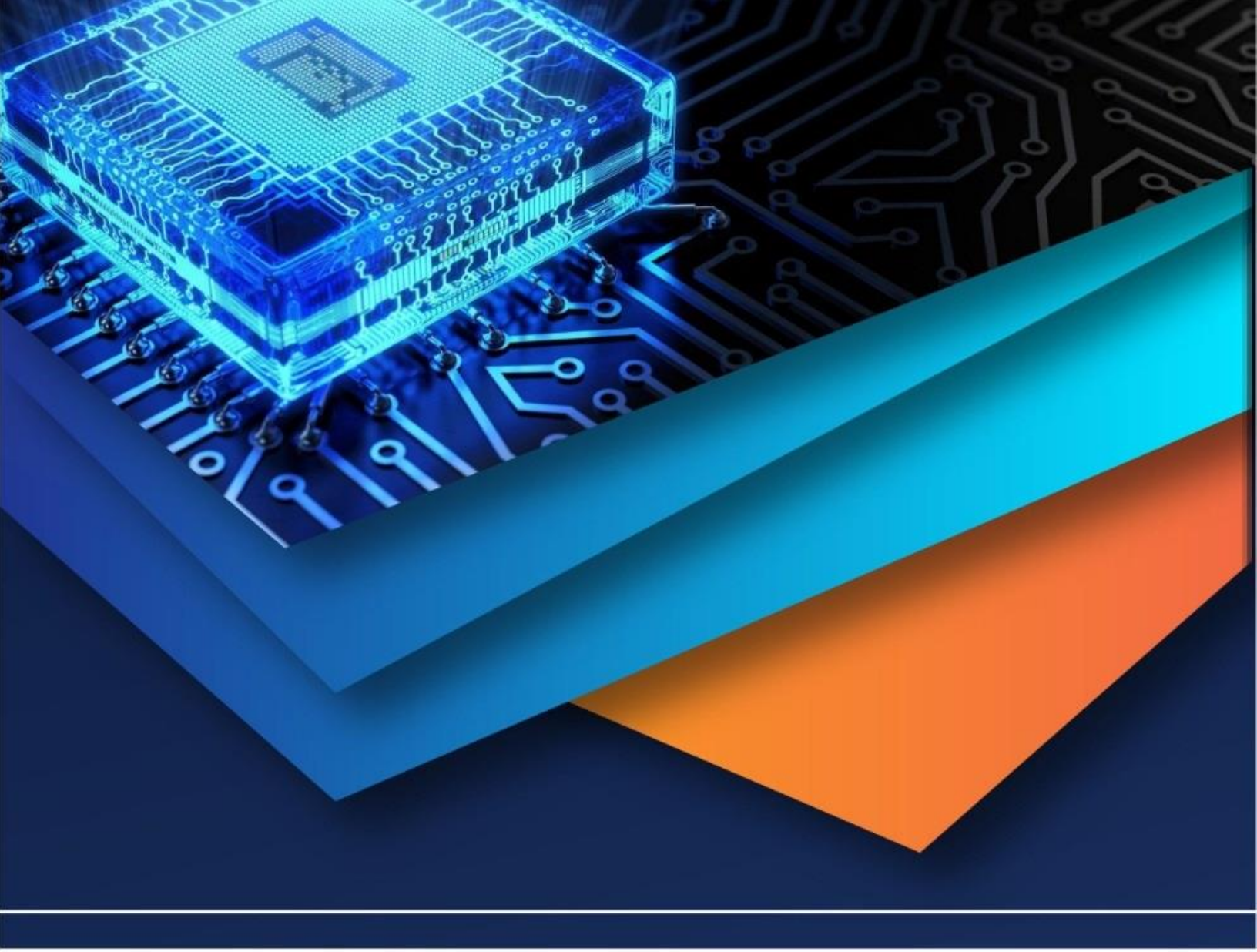
V. CONCLUSION AND FUTURE WORK

A simple and synthetic testbed dataset of 1,500 structured data records was utilized to build & test an interpretable ML system that can predict disease-risks. It was not really applied in a clinical context, although some real clinical assertions were not attempted because this data-set was created for algorithmic purposes. Within this fence the paper has an invulnerable defence position from the methodological point of view. The scheme adopted here did not take into account leakage fields, did not change how the train, validation and test sets are split and compared the classifiers under same input. The highest test accuracy, macro F1 score and ROC-AUC scores were obtained by Gradient Boosting at 0.8178, 0.7652 and 0.9542 respectively. Among the binary risk-screening results, Random Forest (RF) was the best performing model with an accuracy of 0.8978, macro-average F1-score of 0.8738 and ROC-AUC of 0.9632. As depicted in the confusion matrix, the recovery of cardiovascular and no-risk class was greater as there were lesser number of individuals in the diabetes and multiple-risk-cluster classes and the clinical overlap of diabetes and multiple-risk-class. Through feature importance analysis, we found HbA1c, fasting glucose, systolic blood pressure, eGFR, creatinine, LDL cholesterol, BMI, smoking exposure and physical activity to be the most important drivers of our model. These results suggest that there might be clinical and lifestyle type structured disease-risk indicators for controlled disease-risk modelling. Do not diagnose as an impairment. Future investigation of the pipeline should be done with external data acquired in the clinic, checking for calibration drift, fairness test for subgroups, and consideration of changing thresholds to optimize the screening scenario; post-hoc evaluation also with/without clinicians input about the relevance of the features. During the prospective phase the privacy review, institutional approvals, monitoring the changes in the distribution and rules especially for them incorporating with the clinicians monitoring would have to be accomplished as well. The study has immediate impact and the value of the study is how it can be reproduced, its explicit metrics, controlling leakage and the limited interpretation. Well, that's not necessarily bad, just a little added self-control. The absolute lowest benchmark to be considered a credible medical machine learning works. Calibration plots, decision-curve analysis, temporal holdout testing, and a small clinician interface to enable a review of the predictions under practical review conditions and under audit testing and for understanding of the predictions should also be included in the next version.

REFERENCES

[1] G. S. Collins et al., “TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods,” BMJ, vol. 385, Art. no. e078378, 2024, doi: 10.1136/bmj-2023-078378.
[2] K. G. M. Moons et al., “PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods,” BMJ, vol. 388, Art. no. e082505, 2025, doi: 10.1136/bmj-2024-082505.
[3] B. Vasey et al., “Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI,” Nature Medicine, vol. 28, no. 5, pp. 924–933, 2022, doi: 10.1038/s41591-022-01772-9.
[4] A. Zhang, S. Xing, J. Zou, and J. C. Wu, “Shifting machine learning for healthcare from development to deployment and from models to data,” Nature Biomedical Engineering, vol. 6, pp. 1330–1345, 2022, doi: 10.1038/s41551-022-00898-y.
[5] M. Poddar, J. S. Marwaha, W. Yuan, et al., “An operational guide to translational clinical machine learning in academic medical centers,” npj Digital Medicine, vol. 7, Art. no. 129, 2024, doi: 10.1038/s41746-024-01094-9.
[6] N. Pudjihartono, T. Fadason, A. W. Kempa-Liehr, and J. M. O’Sullivan, “A review of feature selection methods for machine learning-based disease risk prediction,” Frontiers in Bioinformatics, vol. 2, Art. no. 927312, 2022, doi: 10.3389/fbinf.2022.927312.
[7] C. L. Andaur Navarro et al., “Risk of bias in studies on prediction models developed using supervised machine learning techniques: Systematic review,” BMJ, vol. 375, Art. no. n2281, 2021, doi: 10.1136/bmj.n2281.
[8] C. L. Andaur Navarro et al., “Completeness of reporting of clinical prediction models developed using supervised machine learning: A systematic review,” BMC Medical Research Methodology, vol. 22, Art. no. 12, 2022, doi: 10.1186/s12874-021-01469-6.
[9] C. L. Andaur Navarro et al., “Systematic review identifies the design and methodological conduct of studies on machine learning-based prediction models,” Journal of Clinical Epidemiology, vol. 154, pp. 8–22, 2023, doi: 10.1016/j.jclinepi.2022.11.015.
[10] O. Efthimiou, M. Seo, K. Chalkou, T. P. A. Debray, M. Egger, and G. Salanti, “Developing clinical prediction models: A step-by-step guide,” BMJ, vol. 386, Art. no. e078276, 2024, doi: 10.1136/bmj-2023-078276.
[11] G. S. Collins et al., “Evaluation of clinical prediction models, part 1: From development to external validation,” BMJ, vol. 384, Art. no. e074819, 2024, doi: 10.1136/bmj-2023-074819.
[12] N. H. Alhumaidi, D. Dermawan, H. F. Kamaruzaman, and N. Alotaiq, “The use of machine learning for analyzing real-world data in disease prediction and management: Systematic review,” JMIR Medical Informatics, vol. 13, Art. no. e68898, 2025, doi: 10.2196/68898.
[13] M. Badawy, N. Ramadan, and H. A. Hefny, “Healthcare predictive analytics using machine learning and deep learning techniques: A survey,” Journal of Electrical Systems and Information Technology, vol. 10, Art. no. 40, 2023, doi: 10.1186/s43067-023-00108-y.
[14] R. Islam, M. Sultana, M. H. Rahman, and M. A. Islam, “A comprehensive review for chronic disease prediction using machine learning algorithms,” Journal of Electrical Systems and Information Technology, vol. 11, Art. no. 27, 2024, doi: 10.1186/s43067-024-00150-4.
[15] D. J. Park, M. W. Park, H. Lee, Y.-J. Kim, Y. Kim, and Y. H. Park, “Development of machine learning model for diagnostic disease prediction based on laboratory tests,” Scientific Reports, vol. 11, Art. no. 7567, 2021, doi: 10.1038/s41598-021-87171-5.
[16] A. Mohamed, M. Abdelrehim, and R. Al-Barazie, “Context matters in machine learning based disease prediction with insights from diverse clinical and symptom data,” Scientific Reports, vol. 15, Art. no. 26855, 2025, doi: 10.1038/s41598-025-26855-8.

[17] N. G. Ramadhan, K. Adiwijaya, W. Maharani, and A. A. Gozali, "Chronic diseases prediction using machine learning with data preprocessing handling: A critical review," IEEE Access, vol. 12, pp. 80698–80730, 2024, doi: 10.1109/ACCESS.2024.3406748.
[18] J. Rashid, S. Batool, J. Kim, M. W. Nisar, A. Hussain, S. Juneja, and R. Kushwaha, "An augmented artificial intelligence approach for chronic diseases prediction," Frontiers in Public Health, vol. 10, Art. no. 860396, 2022, doi: 10.3389/fpubh.2022.860396.
[19] T. Tabashum, R. C. Snyder, M. K. O'Brien, and M. V. Albert, "Machine learning models for Parkinson disease: Systematic review," JMIR Medical Informatics, vol. 12, Art. no. e50117, 2024, doi: 10.2196/50117.
[20] M. B. Makarious, H. L. Leonard, D. Vitale, H. Iwaki, L. Sargent, and A. Dadu et al., "Multi-modality machine learning predicting Parkinson's disease," npj Parkinson's Disease, vol. 8, Art. no. 35, 2022, doi: 10.1038/s41531-022-00288-w.
[21] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn, and M. A. Moni, "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison," Computers in Biology and Medicine, vol. 136, Art. no. 104672, 2021, doi: 10.1016/j.compbiomed.2021.104672.
[22] P. Ghosh, S. Azam, A. Karim, M. Jonkman, A. Anwar, and M. D. R. Islam, "Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques," IEEE Access, vol. 9, pp. 19304–19326, 2021, doi: 10.1109/ACCESS.2021.3053759.
[23] C. M. Bhatt, P. Patel, T. Ghetia, and P. L. Mazzeo, "Effective heart disease prediction using machine learning techniques," Algorithms, vol. 16, no. 2, Art. no. 88, 2023, doi: 10.3390/a16020088.
[24] H. A. Al-Alshaikh, P. Prabu, R. C. Poonia, A. K. J. Saudagar, M. Yadav, H. S. AlSagri, and A. A. AlSanad, "Comprehensive evaluation and performance analysis of machine learning in heart disease prediction," Scientific Reports, vol. 14, Art. no. 7819, 2024, doi: 10.1038/s41598-024-58489-7.
[25] H. El-Sofany, B. Bouallegue, and Y. M. Abd El-Latif, "A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method," Scientific Reports, vol. 14, Art. no. 23277, 2024, doi: 10.1038/s41598-024-74656-2.
[26] X. Liu, W. Zhang, Q. Zhang, C. Chen, T. Zeng, and J. Zhang et al., "Development and validation of a machine learning-augmented algorithm for diabetes screening in community and primary care settings: A population-based study," Frontiers in Endocrinology, vol. 13, Art. no. 1043919, 2022, doi: 10.3389/fendo.2022.1043919.
[27] Y. Jian, M. Pasquier, A. Sagahyroon, and F. Aloul, "A machine learning approach to predicting diabetes complications," Healthcare, vol. 9, no. 12, Art. no. 1712, 2021, doi: 10.3390/healthcare9121712.
[28] O. T. Kee, N. L. Y. Liew, S. W. Lee, and C. S. Thong, "Cardiovascular complications in a diabetes prediction model using machine learning: A systematic review," Cardiovascular Diabetology, vol. 22, Art. no. 13, 2023, doi: 10.1186/s12933-023-01741-7.
[29] M. A. Islam, M. Z. H. Majumder, and M. A. Hussein, "Chronic kidney disease prediction based on machine learning algorithms," Journal of Pathology Informatics, vol. 14, Art. no. 100189, 2023, doi: 10.1016/j.jpi.2023.100189.
[30] H. Khalid, A. Khan, M. Z. Khan, G. Mehmood, and M. S. Qureshi, "Machine learning hybrid model for the prediction of chronic kidney disease," Computational Intelligence and Neuroscience, vol. 2023, Art. no. 9266889, 2023, doi: 10.1155/2023/9266889.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY