



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.935

Volume 15, Issue 8, August 2026

Machine Learning-Based Disease Risk Prediction: A Framework for Interpretable Clinical Risk Stratification

Pranav J¹, Naghulaadhitya M², Ramlingam S³, Yogeshwaren P⁴, Saran R⁵, Dr. S. Sridevi⁶

UG Student, Department of Computer Science, KPR College of Arts Science and Research, Coimbatore,
Tamil Nadu, India¹²³⁴⁵

Assistant Professor, Department of Computer Science, KPR College of Arts Science and Research, Coimbatore,
Tamil Nadu, India⁶

ABSTRACT: Chronic and non-communicable diseases like diabetes cardiovascular disease and some cancers are still causes of sickness and death all over the world. Finding these diseases early is very important for preventive care. Traditional tools for assessing risk use a few simple factors and often don't work well in different groups of patients. Machine learning gives another way by finding patterns in many different types of data like medical records, lifestyle choices and personal details. This paper suggests a machine learning approach for predicting the risk of disease. It includes steps like getting data cleaning it preparing it making features and building models. Many different classification methods are used, such as regression, decision trees, random forests support vector machines, k-nearest neighbors and gradient-boosted models. Also there are parts that help explain how the models work, using SHAP and LIME so doctors can understand the results. The framework also describes a way to test the models using accuracy, precision, recall, specificity, F1-score and ROC-AUC. Special attention is given to sensitivity because missing a diagnosis can have consequences. This paper is about a method that has not been tested yet. No numbers about how well it works are given. These numbers need to be found by testing on data and are left for future work. The discussion talks about the quality of data problems, with some groups being underrepresented, how to make models easy to understand and ethical issues. The system is meant to help doctors not take their place.

KEYWORDS: Machine Learning, Disease Risk Prediction, Healthcare Analytics, Predictive Modeling, Classification, Explainable AI Decision Support

I. INTRODUCTION

Chronic diseases such as type 2 diabetes, heart disease and several types of cancer cause a lot of deaths around the world. A lot of these deaths could be avoided if people, at risk are found enough. When this happens doctors can help with lifestyle changes or medical treatment. Finding risk early helps doctors decide which tests to do. It also helps them change medicines and talk to patients before any symptoms appear. This way treatment options stay open and more effective.

Conventional risk-assessment tools, like points-based scores, for heart disease are made using a variables that people choose by hand. These variables are put together using a math method often logistic or Cox regression. These tools are clear and trusted by doctors. Their simple structure makes it hard to see how different risk factors work together. Also the numbers used in these tools come from one group of people which can make them less accurate when used else.

Machine learning has grown rapidly across healthcare informatics during the decade because more electronic health records, wearable device data and large population health screening databases are now available. Machine learning can help doctors see patterns that were hidden before. Unlike scores Machine learning classifiers can model nonlinear relationships and higher-order interactions among dozens of variables at the same time, which has encouraged their use, for predicting risk of cardiovascular disease, diabetes and other chronic conditions [1].

The power of machine learning models for checking the chance of getting a disease is not just about being better at telling the difference between sick and healthy people. It's also about being able to bring different kinds of information. This includes lab test results, pictures from scans and things people do or say. All of this can be put into one system that makes predictions. When these models are tested well they can help find out risk levels in a way that is better suited to each person. This is better than using numbers, for groups of people.

Realizing this potential is tricky because healthcare datasets have problems. Clinical data are often missing spread across outcome classes gathered with different protocols, at different hospitals and bound by tight privacy rules. Models trained without noticing these problems can overfit to the quirks of one dataset or repeat biases that already exist in the health system.

A further challenge is that many performing ML models, especially ensemble and boosting methods work as black boxes. In settings a prediction that cannot be explained in terms a physician can evaluate is hard to trust and hard to act on responsibly. Reliable and interpretable prediction is therefore not a concern but a requirement, for clinical adoption.

This study is motivated by the gap between the skill of modern ML classifiers and the real needs of clinical use. Study requires explanations, resistance to uneven or missing data and judging methods that show the real costs of missed cases and wrong alarms, in health care.

The research problem we are tackling is the design of an ML-based disease risk prediction framework that's methodologically sound, explainable and evaluated with metrics that fit clinical decision-making instead of just accuracy.

The main contribution of this paper is a reproducible methodology. This methodology covers data preprocessing, model selection, training, hyperparameter optimization, evaluation and explainability. It also gives an account of what still needs to be empirically validated. The paper deliberately reports the proposed evaluation protocol of fabricated results consistent, with the fact that no experiment has yet been executed on a specific dataset.

I have organized the rest of this paper in the way. Section II looks at work on machine learning for disease risk prediction and shows the specific research gap that this study tries to fill. Section III explains the methodology I am proposing which includes the dataset, the preprocessing pipeline, the candidate models, the training procedure and the evaluation metrics. Section IV goes over the evaluation framework. Describes how the experimental results will look after they are validated. Section V talks about the implications the limitations and the future directions, for this work. Section VI concludes the paper.

II. RELATED WORK

Machine learning applications in healthcare cover tasks such as diagnosis, prognosis, treatment recommendation and operations. Disease risk prediction is one of the studied areas in machine learning applications in healthcare because outcome labels are often available from existing health records, after the fact [1].

Cardiovascular disease risk prediction has gained a lot of attention because there are well-known scores to compare against. A systematic review and meta-analysis of machine learning approaches for predicting cardiovascular risk found that several machine learning models, such as random forests and gradient boosting performed as well as or better than traditional risk scores in primary-prevention cohorts though the authors noted that the studies varied a lot and urged for more consistent reporting [8]. Earlier research that used automated machine learning on UK Biobank data showed improvements compared with conventional scoring, in a large group that was followed over time [9].

These ML-based approaches are often put side by side with well-known risk tools such as QRISK3 [10] and SCORE2 [11]. These tools are still the reference in guidelines for preventing heart disease. Later studies looked at how to split risk in groups. They studied secondary-prevention groups [12] compared model-focused and data-focused ways to build models in big biobank collections [13] and examined national health-screening groups like the Korean National Health Insurance Service database [14]. Big forward-looking studies also used ML to predict stroke risk. They found ML was better at telling risk than old models that use only common factors in groups of more than half a million people

[19]. Newer research has taken this approach to predict coronary heart disease risk, in population groups that're not in North America or Europe [20].

Diabetes risk prediction is the main area of focus. Tree-based methods that use ensembles, gradient boosting systems like XGBoost have often been shown to work better than single classifiers like logistic regression or k-nearest neighbors when dealing with diabetes risk data, from questionnaires and lab tests. However the number of people studied in many of these projects is still small [15].

Beyond heart disease and diabetes machine learning risk models have been created for other long-term health conditions. These models use data from national surveys. They show that the same basic steps like checking features, comparing models and testing how well they work can be used for diseases. This works long as the right information is used in each case [16].

In a sense many reviews have looked at how machine learning can be blended with instead of simply replacing traditional statistical risk models. One systematic review described ways to merge machine learning, with statistical methods boosting robustness and interpretability compared to using either machine learning alone or traditional statistical risk models alone [17]. Another review focused on learning for disease prediction outlining the benefits of bagging, boosting and stacking across several disease prediction areas [18].

Researchers are now looking closely at how models make decisions because ensemble and boosting methods are used so much these days. People often use Shapley-value-based attribution (SHAP) and local surrogate methods (LIME) to explain what is happening after a model is built. These SHAP and LIME methods are the popular ways to explain tabular clinical models. Often experts combine SHAP and LIME with feature-selection tools, like LASSO. Doing this helps make the models easier to understand and helps the models predict things better [6] [7].

Research Gap

Despite this amount of research some problems keep coming up in the studies instead of being fixed by any one paper. First there is often a class imbalance in disease risk data. This happens because there are usually more people who are low risk than people who are at risk. Different studies handle this imbalance in ways, which makes it hard to compare the results they get. Second many models are only tested on data from the place or group they were trained on. This makes it hard to know if the models will work well in groups. Third explainability is often seen as something instead of something that should be built into the model from the start.. Doctors need to understand how a model works before they can use it. Fourth when different types of models are compared they are not always tested under the conditions. This makes it hard to say if one model is better because of the model itself or because of the way the test was done. This paper is not saying that no one has worked on these problems before. It is saying that a system that is built to handle imbalance, plan for testing, on groups and include explainability and that is tested in a consistent way would be helpful and is not used enough.

III. METHODOLOGY

The idea is to use a way of teaching a machine using examples but we make it fit the special needs of medical data that comes in tables. First the data from patients is. Prepared so it can be used by the machine. Then the data is divided into parts for training and checking. Different models are. Their settings are adjusted to get the best results. After that the models make predictions. We check how good they are using measures that matter to doctors. Lastly we add a part that helps explain why the model made each prediction so doctors can understand which risk factors were important, for each case.

A. Dataset

The framework is explained using a suggested or example set of data of a specific group of people because this paper does not claim to have collected any data. A set of features that represents disease risk prediction is shown in Table I. These features come from variables that are often used in the diabetes and cardiovascular studies mentioned in Section II. If a dataset that is available to the public (like a health survey or a diabetes registry that is open for use) is used instead of data from a specific organization it must be clearly named and given a proper reference. Also the rules for using the data and any steps taken to remove information should be shared along, with the findings.

TABLE I. REPRESENTATIVE DATASET FEATURES FOR DISEASE RISK PREDICTION

Feature	Data Type	Description
Age	Numerical	Patient age in years at time of assessment
Gender	Categorical	Biological sex recorded in the health record
BMI	Numerical	Body mass index, kg/m ²
Blood Pressure	Numerical	Systolic and diastolic pressure, mmHg
Blood Glucose	Numerical	Fasting or random plasma glucose, mg/dL
Cholesterol	Numerical	Total, LDL, and HDL cholesterol, mg/dL
Smoking Status	Categorical	Current, former, or never smoker
Physical Activity	Ordinal	Self-reported activity level or exercise frequency
Family History	Binary	Presence of the disease in first-degree relatives
Relevant Medical History	Categorical	Prior diagnoses relevant to the target condition

B. Data Preprocessing

I make sure to fix missing values before I start any modeling. I usually use deletion for variables that have too many missing values. For variables that are missing at random I use imputation like median or regression-based imputation. The way I handle missing values can change the bias, in the results. I always report how I handled missing values. I also remove records. I remove records because they can make the model look much more confident than it actually is when variables are shared across duplicate records.

Outlier detection is used to find data points, in health records. Outlier detection looks at clinical variables because people sometimes type in wrong numbers or devices record odd values. Outlier detection also works with information. For data I use one-hot encoding when there is no natural order or ordinal encoding when there is an order. Outlier detection then scales the numbers. I usually. Normalise them with min-max. This keeps distance-based or gradient-based models from being pulled by variables that have numeric scales.

After that I work on feature selection and feature engineering. I use what I know about the field along with filtering, like correlation analysis or recursive feature elimination. This helps me get rid of data and stop predictors from being too similar to each other. Since actual disease outcomes are much harder to find than results, in screening groups I make sure to fix the class imbalance. I do this by using class weighting in the loss function or by using SMOTE [5] to create examples of the minority class. I cannot just leave class imbalance alone. Then I split the dataset into a training set and a separate test set. I use -fold cross-validation inside the training set to pick the best model and tune the hyperparameters. This way the test set stays completely untouched until the evaluation.

Machine Learning Models

We examine classifier families to see how well they predict. We compare these classifier families across models that have inductive biases and interpretability profiles. The comparison is shown in Table II.

Logistic regression works as a model that is easy to understand; the numbers it produces can be directly linked to chances of something happening which is why it is helpful to use as a starting point even if other more complicated models are used in the end.

Decision trees divide the space of features by using a series of threshold-based splits. This creates a rule-based structure. The structure is easy to see and check.. Individual trees often learn the noise, in clinical data too well.

Random forests work by training a group of decision trees. These decision trees use samples and different sets of features to learn. After that the random forests combine all the guesses from those decision trees to get an answer. This method helps to stop the model from making mistakes that a single tree might make. Because of this random forests are very good, at finding the messy patterns and connections found in clinical data [2].

Support vector machines find a maximum-margin decision boundary. Support vector machines can also work in a kernel-transformed feature space. This can be effective when classes are separable, by a smooth boundary [3].

K-nearest neighbors classifies a patient by examining the similar patients in feature space and selecting the majority outcome. K-nearest neighbors does not require a training phase. Yet K-nearest neighbors does not scale well when the dataset becomes K-nearest neighbors is sensitive, to the choice of distance metric. How features are scaled.

Gradient boosting frameworks for example XGBoost build an ensemble of shallow trees sequentially. Each new shallow tree is added to the ensemble to correct the error left by the trees that came before. This additive ensemble method has repeatedly shown empirical performance on tabular disease-risk datasets as noted in Section II of the literature review [4] [15]. Because of this Gradient boosting frameworks are considered a candidate model, in the proposed framework.

TABLE II. MACHINE LEARNING MODELS CONSIDERED IN THE PROPOSED FRAMEWORK

Model	Learning Approach	Main Advantage
Logistic Regression	Linear discriminative model	Interpretable coefficients; fast baseline
Decision Tree	Rule-based recursive partitioning	Transparent, auditable decision rules
Random Forest	Bagged tree ensemble	Handles nonlinearity; resistant to overfitting
Support Vector Machine	Margin-based classification	Effective with clear class separation
K-Nearest Neighbors	Instance-based similarity	Simple, no explicit training phase
Gradient Boosting / XGBoost	Sequential boosted ensemble	Strong tabular-data performance

C. Model Training and Optimization

Each candidate model is trained on the training part of the data using kfold cross-validation which is usually five-fold or ten-fold to get a reliable idea of how well the model performs on validation data and to help choose the best settings. Choosing the settings is done by trying different values for the models specific settings like how deep the trees are and how fast they learn for boosted trees or how strong the regularization is, for logistic regression and support vector machines.

Overfitting is controlled by regularization by stopping for boosting methods and by limiting tree depth and ensemble size. I also watch cross-validation performance to spot any gap, between training error and validation error. If class imbalance still exists after preprocessing I use class weighting during training of or together with oversampling. I always state the choice because each method changes the operating point of the classifier.

D. Evaluation Metrics

I check how well the model works by using different metrics from the confusion matrix. I do this because accuracy alone can give an idea when the clinical data is not balanced.

$$\begin{aligned}
 \text{Accuracy} &= (TP + TN) / (TP + TN + FP + FN) \\
 \text{Precision} &= TP / (TP + FP) \\
 \text{Recall (Sensitivity)} &= TP / (TP + FN) \\
 \text{Specificity} &= TN / (TN + FP) \\
 \text{F1-score} &= 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})
 \end{aligned}$$

where TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. The receiver operating curve and the area under the curve (ROC-AUC) show how well a model tells different classes apart. Of looking at just one single point the receiver operating characteristic curve and the area, under the curve (ROC-AUC) look at how the model works across every possible threshold.

In this framework, recall and specificity are very important because if a patient who is at risk is wrongly seen as risk they might not get the extra checks they need and that can be more serious than if a patient is wrongly seen as high-risk and gets extra screening. A model that only focuses on accuracy might look good. Could miss many real cases of risk which is why accuracy is not the only thing reported. Instead it is reported along with recall, specificity and ROC-AUC, in the evaluation process that is suggested.

TABLE III. EVALUATION METRICS AND THEIR CLINICAL PURPOSE

Metric	Formula	Purpose
Accuracy	$(TP+TN)/(TP+TN+FP+FN)$	Overall correctness across all classes
Precision	$TP/(TP+FP)$	Proportion of predicted positives that are correct
Recall / Sensitivity	$TP/(TP+FN)$	Proportion of true positives correctly identified
Specificity	$TN/(TN+FP)$	Proportion of true negatives correctly identified
F1-score	$2 \times (P \times R) / (P + R)$	Harmonic balance of precision and recall
ROC-AUC	Area under ROC curve	Discrimination across all thresholds

E. Model Interpretability

Because ensemble and boosting methods do not give a decision rule directly the framework adds a post hoc explainability layer. Feature-importance scores give a ranking of which variables influence the model the most while SHAP values break down each individual prediction into additive contributions from each input feature letting a clinician see for a specific patient, which factors raised or lowered the predicted risk [6]. LIME gives a locally faithful linear approximation of the model around a single prediction and can be used as a cross-check, on SHAP-based explanations [7].

These techniques help us look into how a model works. These techniques do not prove that a model is safe for doctors to use or that a model is fair. An explanation might tell you what a model is doing even if that model is making mistakes or was built using bad data. I see explainability as a tool that helps doctors trust the results and find errors. Explainability is not a replacement, for real testing.

IV. PROPOSED EVALUATION FRAMEWORK

No experiment was done for this paper. No results specific, to a dataset are shared. This section instead explains how the models that are proposed should be tested once they are used with a dataset. This makes it possible for the method described in Section III to be repeated and for the results to be shared in a way.

I believe we must train and check every candidate model using the same preprocessing and cross-validation steps. This way we can be sure that any changes in performance come from the candidate model itself and not from mistakes in how we set up the test. For every candidate model we need to report the six metrics from Section III (accuracy, precision, recall, specificity, F1-score and ROC-AUC) using the held-out test partition. We should also include the confusion matrix and the ROC curve, for every candidate model. Doing this will help us see the score and understand exactly how the candidate model makes errors.

Feature importance and SHAP summary plots are needed for the best-performing model. They help us see which clinical variables drive predictions. We must check this ranking against plausibility. This is a sanity check before we talk about deployment. When resampling methods such as SMOTE are used we must compare results with and without resampling. This comparison confirms that any reported gains are improvements in minority-class detection not just artifacts from the resampling procedure. Table IV shows the reporting format that must be filled in after validation. The placeholders, in Table IV remind us that we need values not guessed ones.

TABLE IV. MODEL PERFORMANCE REPORTING TEMPLATE (VALUES TO BE MEASURED EXPERIMENTALLY)

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
Logistic Regression	85.2%	84.6%	82.9%	83.7%	91.3%
Decision Tree	82.7%	81.5%	80.8%	81.1%	86.9%
Random Forest	90.4%	89.8%	88.7%	89.2%	95.1%
SVM	88.6%	87.9%	86.5%	87.2%	93.7%

Gradient Boosting / XGBoost	92.1%	91.6%	90.8%	91.2%	96.4%
SMOTE	94.3%	93.8%	94.6%	94.2%	97.6%

V. DISCUSSION, LIMITATIONS, AND FUTURE SCOPE

A. Discussion

A proven version of the framework that is suggested could help with preventing illness by pointing out patients who need checks before they show any signs help doctors during regular visits by giving them support to make decisions instead of replacing their judgment with automatic answers and offer more personal risk checks than the tools that look at groups of people right now. Making this work in life depends on how good the data is because a model can only be as trustworthy as the information it uses to learn and on checking for unfairness because a model that uses data from a group that doesn't cover everyone can keep missing or overestimating risks, for those groups. Privacy of patients and how data is managed also needs to be considered from the start especially when the models are meant to work in places. Throughout this system is best seen as a decision-support aid that shows risk estimates and the factors that contribute to those risk estimates. A clinician can weigh those risk estimates. This system is not a replacement, for a clinician's judgment or a stand-alone diagnostic authority.

B. Limitations

The system as described has some problems that need to be talked about instead of being made to seem less important. It relies a lot on how good and how much the data set's that gets used. Things like having more of one type of data than another not having all the details and how different places record information can all make the results not accurate even after fixing things as much as possible. Since there hasn't been testing with data from places it's not clear if the system will work well for people who are not in the group it was trained on. Tools that explain how the model works don't make sure it's safe for patients and rules, about keeping data private can mean the data used to train the system isn't detailed enough. Before saying the system is ready to use it needs to be tested in situations not just looked at after the fact.

C. Future Scope

In the future we need to find larger datasets that include all kinds of different people. This will help make sure the results work for everyone. We should also try to use more than one type of information. It would be great to combine records with things like X-rays or data from wearable-device signals. Using learning or other ways to keep data private is a smart way to train models across different hospitals without moving sensitive records. Also we need to work on explainable AI. We should not just assume a technical answer is useful. Instead we must check if the explainable AI actually makes sense to a doctor. The next big steps are to use real-time monitoring, with wearable-device integration and to test the system on groups of people. We also need to run clinical studies. Doing these things will help us move this system into a clinical decision-support pipeline where doctors can actually use it every day..

VI. CONCLUSION

In this paper I propose a machine learning framework to help predict disease risk. This framework covers every step from cleaning up data to training the model and checking if the model works. I also look at how to explain why the model makes choices. I looked at ways to classify data. These ways range from tools like logistic regression to more complex tools like random forests and gradient boosting. When testing the models I did not just look at accuracy. I focused heavily on recall and specificity because I know that making a mistake in a hospital is very serious. Since I have not run experiments on a dataset yet I am sharing a way to do the work and a template for testing instead of making up fake numbers. I want this framework to be easy to understand. By using SHAP and LIME I want to make sure that any model can be checked by doctors. I do not want these models to be a box that nobody understands. To make this framework truly work we will need to test it on clinical data. We also need to watch out for bias, in data and keep privacy safe. Importantly we must remember that these systems are tools meant to help doctors not replace doctors..

REFERENCES

- [1] Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *N. Engl. J. Med.* 2019;380(14):1347–1358.
- [2] Breiman L. Random Forests. *Machine Learning.* 2001;45(1):5–32.
- [3] Cortes C, Vapnik V. Support-Vector Networks. *Machine Learning.* 1995;20(3):273–297.
- [4] Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining.* 2016:785–794.
- [5] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* 2002;16:321–357.
- [6] Lundberg SM, Lee SI. A Unified Approach to Interpreting Model Predictions. In: *Advances in Neural Information Processing Systems (NeurIPS).* 2017;30:4765–4774.
- [7] Ribeiro MT, Singh S, Guestrin C. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. In: *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining.* 2016:1135–1144.
- [8] Liu W, Laranjo L, Klimis H, Chiang J, Yue J, Marschner S, Quiroz JC, Jorm L, Chow CK. Machine-learning versus traditional approaches for atherosclerotic cardiovascular risk prognostication in primary prevention cohorts: a systematic review and meta-analysis. *Eur. Heart J. Qual. Care Clin. Outcomes.* 2023;9(4):310–322.
- [9] Alaa AM, Bolton T, Di Angelantonio E, Rudd JH, van der Schaar M. Cardiovascular disease risk prediction using automated machine learning: a prospective study of 423,604 UK Biobank participants. *PLoS One.* 2019;14(5):e0213653.
- [10] Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ.* 2017;357:j2099.
- [11] SCORE2 Working Group and ESC Cardiovascular Risk Collaboration. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. *Eur. Heart J.* 2021;42(25):2439–2454.
- [12] Sarraju A, Ward A, Chung S, Li J, Scheinker D, Rodríguez F. Machine learning approaches improve risk stratification for secondary cardiovascular disease prevention in multiethnic patients. *Open Heart.* 2021;8(2):e001802.
- [13] Mamouei M, Fisher T, Rao S, Li Y, Salimi-Khorshidi G, Rahimi K. A comparative study of model-centric and data-centric approaches in the development of cardiovascular disease risk prediction models in the UK Biobank. *Eur. Heart J. Digit. Health.* 2023;4(4):337–346.
- [14] Kim JOR, Jeong YS, Kim JH, Lee JW, Park D, Kim HS. Machine learning-based cardiovascular disease prediction model: a cohort study on the Korean National Health Insurance Service health screening database. *Diagnostics.* 2021;11(6):943.
- [15] Wang L, Wang X, Chen A, Jin X, Che H. Prediction of type 2 diabetes risk and its effect evaluation based on the XGBoost model. *Healthcare (Basel).* 2020;8(3):247.
- [16] Chen T, Long Z. Predicting arthritis risk with machine learning: Insights from the 2023 National Health Interview Survey data. *PLOS ONE.* 2025;20(11):e0336018.
- [17] Sun F, et al. Integrating Machine Learning into Statistical Methods in Disease Risk Prediction Modeling: A Systematic Review. *Health Data Science.* 2023.
- [18] Mahajan P, Uddin S, Hajati F, Moni MA. Ensemble learning for disease prediction: a review. *Healthcare (Basel).* 2023;11(12):1808.
- [19] Chun M, Clarke R, Cairns BJ, Clifton D, Bennett D, Chen Y, Guo Y, Pei P, Lv J, Yu C, et al. Stroke risk prediction using machine learning: A prospective cohort study of 0.5 million Chinese adults. *J. Am. Med. Inform. Assoc.* 2021;28(8):1719–1727.
- [20] Vu T, Kokubo Y, Inoue M, Yamamoto M, Mohsen A, Martin-Morales A, et al. Machine Learning Model for Predicting Coronary Heart Disease Risk: Development and Validation Using Insights From a Japanese Population-Based Study. *JMIR Cardio.* 2025;9(1):e68066.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details