

Original Research

Received 28 May 2026

Revised 08 July 2026

Accepted 22 July 2026

Natural Resources for Human Health 2026, Vol. 6 (13s): 1358–1365

KEYWORDS:

Logistic Regression (LR), Random Forest (RF), Decision Tree (DT), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), XGBoost (XGB), Gradient Boosting (GB), Gaussian Naïve Bayes (GNB), AdaBoost (AB), Machine Learning (ML), Artificial Intelligence (AI), Deep Learning (DL), Precision (Pr), Recall (Rc), F1-Score (F1), Accuracy (Ac).

Stacking Model for Diabetes, Heart, Liver and Kidney Disease Prediction

Ramdas Pandurang Bagawade¹, Dr. Thirupurasundari D.R.²

¹ Bharath Institute of Higher Education and Research, India (Chennai)

² Bharath Institute of Higher Education and Research, India (Chennai)

ABSTRACT: Predicting chronic diseases, such as diabetes, cardiovascular diseases, and organ (liver and kidney) disease, is becoming very important in health care as it allows for early intervention and detection of diseases. It can also save lives and reduce healthcare costs to patients. Due to the combined capabilities of several algorithms (machine learning (ML) algorithms), this study examines the use of Random Forest (RF), Decision Tree (DT), and Logistic Regression (LR) alongside Gradient Boosting (GB) algorithms. The technique stacking combines different models of the algorithms and harnesses their capabilities in making a prediction and then improves classification. The preliminary results show that in respective prediction of diabetes, cardiovascular disease, and organic (liver and kidney) disease, the results of the stacking model are better in all of prediction accuracy, prediction sensitivity and prediction specificity, and greater than all other individual algorithms across all diseases.

1. INTRODUCTION

Predictive healthcare, especially with the use of machine learning, can help develop novel intervention methods to better manage health care systems. Due to the complexities of data in medicine and the need for personalized medicine, the techniques of machine learning are essential and very useful. Various research studies have demonstrated that individual machine learning approaches (RF, DT, LR, GB) can achieve significant results in disease states prediction in different healthcare systems. Nonetheless, their independent functioning may be affected by the core algorithmic limitations like overfitting, poor generalization, and the inability to learn complex data relationships. As a remedy for these issues, stacking as a type of ensemble learning, integrates diverse algorithms to enhance predictive power and strengthen the overall predictive model. The primary goal of this study is to design a stacking ensemble model to predict a patient's diabetes, heart, liver, and kidney disease utilizing the combined powers of RF, LR, DT, and GB. The aim of the planned model is to add value to disease classification and empower healthcare analytics. In this study the authors describe their strategy, and the predictive healthcare technologies and medical diagnostics that this approach is likely to impact.

The use of ensemble based approaches, stacking, and to some extent, other methods, has gained a lot of popularity in healthcare analytics both for the improvement of predictive and overall robustness. Stacking also integrates the strength of four or more diverse ML algorithms to use the power of several predictive models that form the base learning. This has been extensively researched, especially for the prediction of chronic conditions of diabetes, heart, liver and kidney diseases. The efficacy of disease prediction algorithms has been described in the literature for disease prediction using RF, LR, DT, and GB. RF is especially noted for its superior ability to, among other things, effectively use high-dimension data and reduce the likelihood of the model overfitting through the use of bagging, while LR has the ability to effectively handle simple binary decisions and provides interpretable models.

DT is simple and good as a stand-alone algorithm, and GB increases accuracy by iteratively reducing weak learners. Yet, multiple drawbacks, typically the inability to generalize, the overfitting of models, and the inability to model the inherent data complexities, are issues within medical datasets, resulting in poor performance from most stand-alone algorithms. Most other ensemble algorithms, including bagging and boosting, struggle significantly. Therefore, research indicates that stacking algorithms outperform the rest within the majority of medical applications. For instance, in heart disease classification,

stacking algorithms perform better based on metrics such as balanced accuracy, precision, and recall than any other algorithms [19]. With regards to kidney disease classification, studies have emphasized the benefits that stacking models offer in the continuum of the classification performance on imbalanced datasets [20]. With a single framework that encompasses the stacking algorithm's capabilities, the reliability of any disease prediction system will greatly improve, thereby mitigating most of the issues caused by the use of a single disease prediction framework.

Few have attempted to apply stacking multi-class chronic disease predictions algorithm. With the majority of the studies spanning over a concentrated disease focus, the attempts to employ stacking algorithm in disease prediction have failed sufficiently to address the need of incorporating chronic disease spectrum. This study intends to combine the novel usage of stacking framework in multi-class predictions along with the improvement of prediction systems to the spectrum of Chronic Diseases, namely, Diabetes, Heart Disease, Kidney, and Liver Disease, by integrating the RF, LR, DT, and GB algorithms in a single framework. The use of divergent chronic disease predictors is geared toward elevating the ensemble application in healthcare systems. Overtime, diabetes can result in various additional health issues, such as cardiovascular disease (risk of heart attack), nerve damage, kidney disease, eye damage, and foot complications [30]. These issues stem from the harm inflicted by persistently elevated blood sugar levels on blood vessels and nerves across the body.

2. LITERATURE REVIEW

2.1 Paper [18]:

In this paper author discussed about three main classes of ensemble techniques (Bagging, Boosting and Stacking) used in machine learning.

2.2 Paper [21]:

The stacking ensemble method enhances the model's accuracy. It can be observed that the stacking ensemble approach results in increased accuracy for the random forest, achieving 99.02%, while there is a slight reduction in the accuracy of both SVM and KNN.

2.3 Paper [22]:

The Logistic Regression model registered predictive accuracy of 86.67%. The K-Nearest Neighbor method achieved accuracy of 88.33%. The Naïve Bayes model gave 86.66% accuracy. The accuracy of the ensemble of these three classifiers added up to 90% with regard to the multiple assessment metrics utilized to evaluate model performance.

2.4 Paper [23]: Utilizing Stacking Model Fusion together led to an accuracy boost of 90.46%. The research also highlights the importance of choosing a sampling strategy that aligns with the characteristics of the dataset

2.5 Paper [24]:

In the initial layer, KNN, NB, and GBDT are selected as the foundational classifiers, and 5-fold cross-validation is utilized to mitigate the risk of over-fitting. A basic logistic regression model serves as the meta-classifier in the subsequent layer, and it is observed that the ultimate model reaches greater accuracy than the individual classifiers, showcasing strong generalization capability.

2.6 Paper [25]:

Seventy-five percent of the dataset was used for training and the remaining 25% for testing. The parametric metrics such as the F1 score, recall, as well as AUC and accuracy, suggested that ensemble stacking reaches a maximum of 85.6, 86.5, 88.1, and 98.4 percent, respectively. Hence, ensemble stacking tends to be the most optimal learning algorithm.

2.7 Paper [26]:

This study evaluated stacked ensemble learning for heart disease prediction in comparison to traditional machine learning methods in the case of binary and multiclass predictions. Results show that the stacked ensemble model exceeded all individual models on multiple metrics such as accuracy, precision, recall, F1, and AUC, especially for binary case predictions.

2.8 Paper [27]:

The model utilizes five different classifiers: Support Vector Machine, Naïve Bayes, K-Nearest Neighbour, XGB, and Random Forest to enhance prediction accuracy. For the model analysis, various parameters such as accuracy, recall, precision, and F1 score are utilized. The experimental findings of the suggested ensemble model are contrasted with traditional machine learning models. The analysis indicates that the proposed stacking ensemble framework attained an improved accuracy of 90.16%.

3. METHODOLOGY

1. Data Collection and Pre-processing

Used dataset of diabetes [33, 34], heart [31, 32], liver [36], and kidney [35] diseases from kaggle open access website. Then applied different data cleaning techniques for recovering missing values and removing incomplete rows and outliers. Then performed feature selection on the available dataset to identify the most relevant attributes for each disease, ensuring that redundant or irrelevant features are excluded. At last available dataset is divided randomly into 70 and 30 percentage (e.g., 70% training, 15% validation, 15% testing) to ensure unbiased evaluation.

2. Selection of Base Model

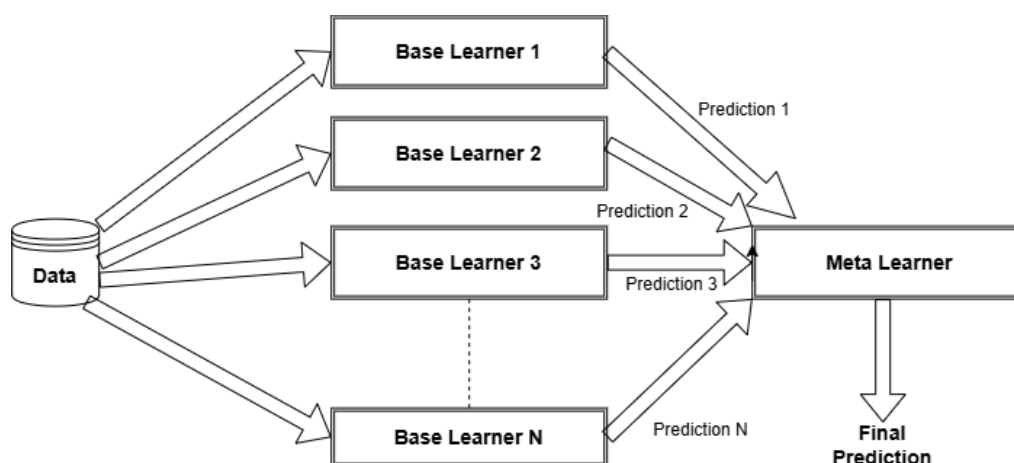
The base models that are chosen are LR, RF, SVM, KNN, AB, GB, and DT.

3. Implementation of Base Models

The constituent models (LR, RF, SVM, KNN, AB, GB, and DT) are learned on the training data independently, and the results are evaluated.

4. Stacking Ensemble Model Construction

First, Level-0 Models: Use LR, RF, SVM, KNN, AB, GB, and DT as base learners. Each model independently generates predictions on the dataset. Meta-Model (Level-1): Utilize a higher-level algorithm (e.g., Logistic Regression) to aggregate predictions from the base learners. The input to the meta-model comprises the outputs (predicted probabilities) from the base learners. Apply cross-validation during the training of base learners and meta-model to ensure robustness and avoid over-fitting. Final prediction is produced by Meta model as shown in fig1. The whole process is depicted in fig2.



Stacking Approach in Ensemble Learning

Fig 1: Stacking Approach in Ensemble Learning [27]

5. Model Evaluation

Evaluated the stacking ensemble model on the test dataset using the same performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Compared the results of the stacking model against individual base models to assess

performance improvements in accuracy, sensitivity, and specificity. Then conducted statistical significance tests (e.g., paired t-tests) to validate the enhanced performance of the stacking model.

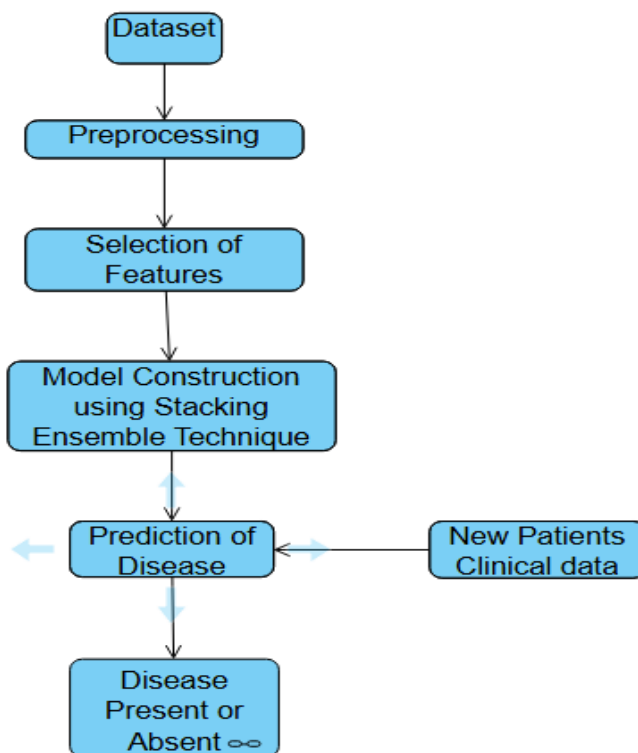


Fig 2: Stacking Approach in Ensemble Learning [19]

3. RESULTS AND DISCUSSION

Table 1. Results obtained for Heart1 Dieses (using Indian Dataset [29])

Model	Ac	Pr	Re	F1
LR	0.6936	0.6494	0.8441	0.734
DT	0.7367	0.7642	0.686	0.723
RF	0.723	0.6942	0.7989	0.7429
KNN	0.703	0.7	0.7122	0.7061
SVM	0.7194	0.7654	0.6341	0.6935
GB	0.7378	0.7528	0.7095	0.7305
AB	0.7258	0.7354	0.7069	0.7209
Stacked Model with 7	0.7145	0.724	0.695	0.7092
Stacked Model with 4	0.7371	0.7499	0.7131	0.731

Table 2. Results obtained for Heart2 Dieses (using Foreign Dataset [30])

Model	Ac	Pr	Re	F1
LR	0.8049	0.7692	0.8738	0.8182
DT	0.9854	1	0.9709	0.9852
RF	0.9854	1	0.9709	0.9852
KNN	0.8341	0.8	0.8932	0.844

SVM	0.8879	0.8509	0.9417	0.894
GB	0.9171	0.9135	0.9223	0.9179
AB	0.878	0.89	0.8641	0.876
Stacked Model with 7	0.9122	0.9126	0.9126	0.9126
Stacked Model with 4	0.8976	0.8796	0.9223	0.9005

Table 3. Results for Diabetes1 Indian Dataset [31]

Model	Ac	Pr	Re	F1
LR	0.0888	0.0888	1	0.1631
DT	0.9437	0.6447	0.8148	0.719
RF	0.9753	0.9815	0.7361	0.8413
KNN	0.0888	0.0888	1	0.1631
SVM	0.9112	0	0	0
GB	0.9757	0.9937	0.7315	0.8427
AB	0.9737	0.9368	0.7546	0.8359
Stacked Model with 7	0.9733	0.9218	0.7546	0.8359
Stacked Model with 4	0.9679	0.852	0.7731	0.8427

Table 4. Results for Diabetes2 Foreign Dataset [32]

Model	Ac	Pr	Re	F1
LR	0.7532	0.6491	0.6727	0.6607
DT	0.7597	0.6552	0.6909	0.6726
RF	0.7208	0.6034	0.6364	0.6195
KNN	0.6883	0.5745	0.4909	0.5294
SVM	0.7273	0.6327	0.5636	0.5962
GB	0.7468	0.6379	0.6727	0.6549
AB	0.7338	0.625	0.6364	0.6306
Stacked Model with 7	0.7597	0.6552	0.6909	0.6726
Stacked Model with 4	0.7208	0.6	0.6545	0.6261

Table 5. Results obtained for Kidney (using Dataset [33])

Model	Ac	Pr	Re	F1
LR	0.975	1	0.9615	0.9804
DT	0.975	1	0.9615	0.9804
RF	0.9875	1	0.9808	0.9903
KNN	0.6875	0.8857	0.5962	0.712
SVM	0.65	0.65	1	0.7879
GB	1	1	1	1
AB	0.9875	1	0.9808	0.9903
Stacked Model with 7	0.9875	1	0.9808	0.9903
Stacked Model with 4	1	1	1	1

Table 6. Results obtained for Liver (using Dataset [34])

Model	Ac	Pr	Re	F1
LR	0.6552	0.6486	0.9863	0.7826
DT	0.6121	0.6628	0.7808	0.717
RF	0.6034	0.6337	0.8767	0.7356
KNN	0.6121	0.6628	0.7808	0.717
SVM	0.6293	0.6293	1	0.7725
GB	0.6897	0.7033	0.8767	0.7805
AB	0.6552	0.6941	0.8082	0.7468
Stacked Model with 7	0.6466	0.6818	0.8219	0.7453
Stacked Model with 4	0.6121	0.6628	0.7808	0.717

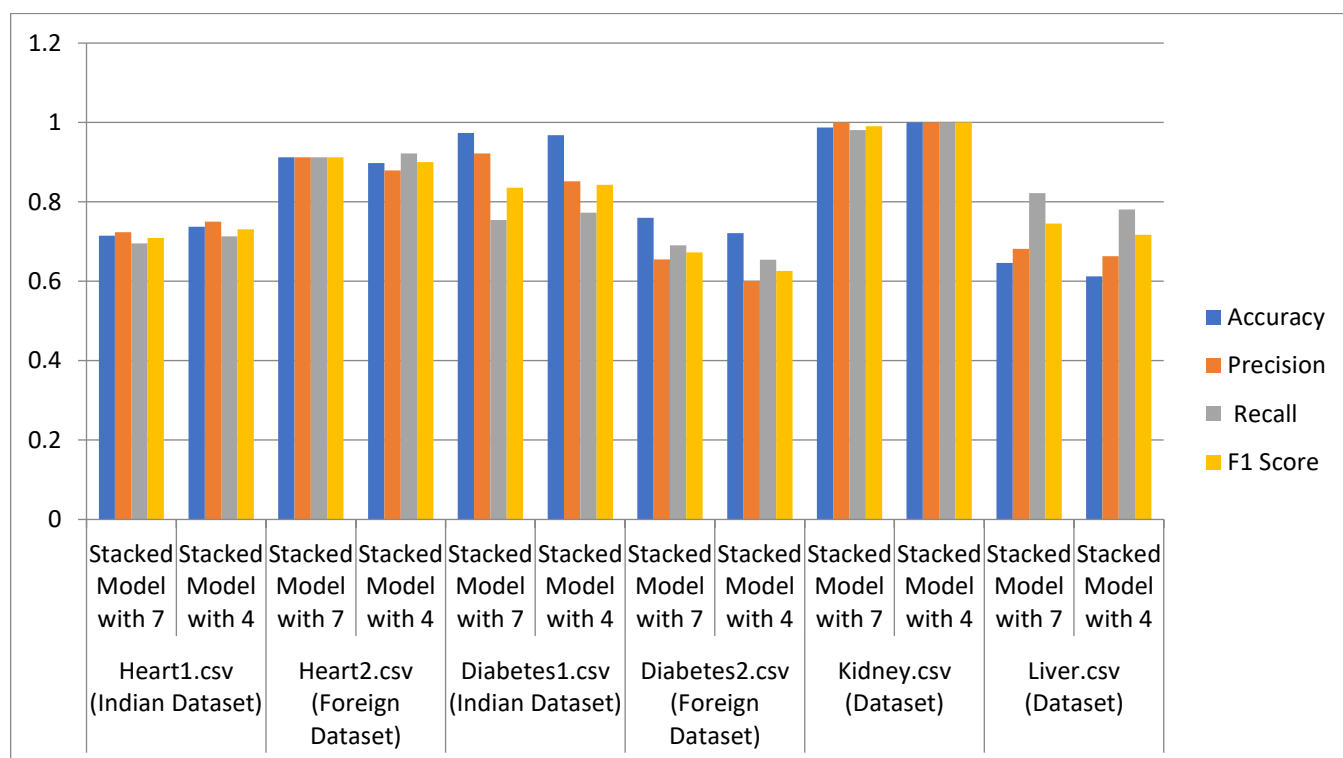


Fig 3. Performance of Stacked Model.

Table-1 and Fig3 shows the performance of stacked model, where first time 7 base models (LR, RF, SVM, KNN, AB, GB, and DT) performance is given to LRmeta model to get the final result as show in fig1. Second time 4 base models (LR, RF, GB, and DT)performance is given to LRmeta model to get the final result as show in fig1. For this result dataset utilized from kaggle website.

4. CONCLUSION

This research attempts to predict chronic diseases, including diabetes, heart disease, liver disease, and kidney disease, through those stack ensemble models. Stacking of RF, LR, DT, and GB offers those improvements in the proposed stacking method, and as such, is seen to be of greater value and more accurate in the diabetes case, as seen in the study. It exhibits that the stacking method models are improved in concerns of over-fitting, lack of generalization, and divergence of interests, as portrayed by the absence of constituent models.

The stacking method integrates multiple divisions of predictive potential, making it invaluable in the field of healthcare analytics. It goes well beyond disease prediction to the realms of predictive analytics, personalized patient care, and clinical decision-making. This research lays the foundation for the stacking ensemble models, particularly in the area of predictive patient care, and guides the balance of accuracy against the need for more scalable and reliable healthcare solutions in fast changing healthcare systems. Performance of Staking model is better compare to performance of base models as shown in Fig3. Ultimately, the stacking method is an excellent model for machine-learning healthcare models. Stacking method is not a destination, but it is a promising, available path toward more efficient and effective patient care systems.

ACKNOWLEDGEMENTS

The authors acknowledge the dataset provided by the <https://www.kaggle.com/>.

AUTHOR CONTRIBUTIONS

Ramdas Pandurang Bagawade: Conceptualization, Investigation, Formal Analysis, Laboratory Analysis, Methodology.

Dr. Thirupurasundari D.R: Supervision, Writing—Review & Editing.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

ETHICS STATEMENT

Ethical approval was not required for this study.

DATA AVAILABILITY STATEMENT

Data supporting the findings of this study are publically available from the kaggle website.

REFERENCES

- [1] https://infyspringboard.onwingspan.com/web/en/app/toc/lex_auth_0130944105066496001885_shared/overview
- [2] <https://viso.ai/deep-learning/ensemble-learning/>
- [3] <https://www.geeksforgeeks.org/gradientboosting-vs-adaboost-vs-xgboost-vs-catboost-vs-lightgbm/>
- [4] <https://raw.githubusercontent.com/mwaskom/seaborn-data/master/penguins.csv>
- [5] <https://www.geeksforgeeks.org/information-gain-and-mutual-information-for-machine-learning/>
- [6] Zhi-Hua Zhou, “Ensemble Methods Foundation and Algorithm”, CRC press: Taylor and Francis Group, (2012).
- [7] J.Huang, Z.-H. Zhou, H.-J. Zhang, and T. Chen, “Poseinvariant face recognition”. In Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, pp: 245–250, Grenoble, France, (2000).
- [8] Giacinto, F. Roli, and L. Didaci, “Fusion of multiple classifiers for intrusion detection in computer networks”. In: Pattern Recognition Letters, pp: 1795–1803, (2003).
- [9] Hansen L, Solomon P, “Neural Network Ensembles”. In: IEEE transactions on pattern analysis, pp: 993-1001, (1990).
- [10] Thomas. G. Dietterich, “Ensembles methods in machine learning.” In: International workshop on multiple classifiers systems. pp: 1-15, Springer, (2000).
- [11] Blum A, Rivest R.L, Training a 3-node neural network is NPComplete. In: Proceedings of the 1988 workshop on computational learning theory, pp: 9-18, San Francisco, CA. Morgan Kaufmann, (1988).
- [12] R.P.Bagawade, R Manda, J. Gavit, V. Karanjakar, A. Lohar “Heart Disease Prediction Using Deep Learning”. IRJMETs, Vol. 06, Issue.5, pp: 8186-8191, May 2024.
- [13] R.P.Bagawade, D.R Thirupurasundari, “A Review Data Pre-Processing Techniques in Machine Learning”. In Proceedings of Hinweis Second International Conference , pp: 3937-3946, RTET, (2024), <https://www.scopus.com/pages/publications/105034156803>

- [14] R.P.Bagawade, D.R Thirupurasundari, "Disease Prediction Using Machine Learning". JETM , Vol. 74, no.1, pp: 1-9, Dec 2024
- [15] R.P.Bagawade, D.R Thirupurasundari, "A Review of Feature Selection Techniques for Machine Learning Algorithms used for Disease Prediction". In Proceedings of ICRTSET Fifth International Conference , pp: 330-335, (2025).
- [16] Ramdas Bagawde, Sanket Pehera, Shubham Gadekar, Utkarsh Kulkarni, "Liver, Kidney, Heart and Diabetes Disease Prediction using Machine Learning". International Journal of Innovative Research in Computer and Communication Engineering , Vol. 13, Issue 03, pp: 2728-2734, March 2025
- [17] R.P.Bagawade, D.R Thirupurasundari, "Ensemble Machine Learning Techniques". JJETA , Vol. 15, Issue.4, pp: 15-23, Apr 2025
- [18] E.G. Kumar, Dr. Madan Lal Saini, Syed Abbas Khadar Ali and Beeram Bhanu Teja, " A Clinical Support System for Prediction of HeartDisease using Ensemble Learning Techniques", ICSCNA 2023
- [19] Nazim Nisar Itoo, Vijay Kumar Garg, "Heart Disease Prediction using a Stacked Ensemble of Supervised Machine Learning Classifiers", MECON 2022
- [20] Siti Juwariyah, Ema Utami, "Enhancing Stacking Model Fusion for Heart Disease Prediction", TIME-E August 2024
- [21] Pan Liu, Li LiYihao Li, "A Stacking-Based Heart Disease Classification Prediction Model", DSA 2023
- [22] D. Rosy Salomi Victoria, Ramalakshmi N R, Mayuri Ki,Meghna Ganesh Kumar "An Ensemble Stacking Methodology forIdentifying Heart Disease", ICOSEC-2023
- [23] Amirthaa M, Anitha Julian, Gerardine Immaculate Mary, Anto Lourdu Xavier Raj Selvarathinam"Improving Heart Disease Prediction with Stacked Ensemble Learning: A Comparison of Binary and Multiclass", InCoWoCo2024
- [24] Veena Kumari H M, Suresh D S "Performance Analysis of Base and MetaClassifiers and the Prediction of CardiovascularDisease using Ensemble Stacking", IDCIoT 2023
- [25] Veena Kumari H M, Suresh D S "Thyroid Disease Prediction Model on BoostingbasedStacking Ensemble Approach", I2CT 2023
- [26] <https://viso.ai/deep-learning/ensemble-learning/>
- [27] <https://medlineplus.gov/diabetescomplications.html>
- [28] R.P.Bagawade, D.R Thirupurasundari, "Heart, Liver, Lungs, Diabetes and Kidney Disease Prediction using Machine Learning and Voting Ensemble Technique", ICSFT 2026 International IEEE Conference 24-25 April 2026, DOI: 10.1109/ICSFT67370.2026.11566293
- [29] Heart Disease Dataset," Kaggle, [Online] Available: <https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset>
- [30] Cardiovascular Disease dataset," Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset>
- [31] "Diabetes Dataset," Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/mathchi/diabetes-data-set/data>
- [32] "Diabetes prediction dataset," Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/iammustafatz/diabetes-prediction-dataset/data>
- [33] "Chronic KIdney Disease dataset," Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/mansoordaku/ckdisease>
- [34] "Indian Liver Patient Records," Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/uciml/indian-liver-patient-records/data>