A MULTI-ORGAN MACHINE LEARNING FRAMEWORK FOR EARLY ASSESSMENT OF TREATMENT VERSUS TRANSPLANT NECESSITY

Asst. Prof. R.P. Bagawade, Ms. A.P. Sawant, Mr. A.R. Mahadik, Mr. P. C. Kalsait, Mr. O.V. Bansode

Department of Computer Engineering, Government College of Engineering and Research, Avasari Khurd. Savitribai Phule Pune University, Pune

ram.vish22@gmail.com, ankitasawant8998@gmail.com, ayushmahadik49@gmail.com, pradhumnyakalsait@gmail.com, onkarbansode77@gmail.com

Abstract

This study presents a decision support model based on machine learning that predicts whether a patient is likely to re- quire standard medical care, or if organ transplantation is necessary. The model focuses on the kidneys, liver, heart, and lungs. It helps in improving patient outcomes and reducing clinical decision-making delays by the early identification of severe organ dysfunction. This model uses several machine learning algorithms independently also along with ensemble techniques, which trained for each organ based on the relevant clinical and diagnostic variables. Each model is assessed based on accuracy, precision, recall, and F1-score in order to quantify organ specific models. The goal of the described method is to provide clinicians with a data driven, early indication of the severity of the patient's condition, in order to facilitate timely medical intervention and optimized resource allocation.

Keywords: *Machine learning clinical decision support systems, health- care analytics.*

1 INTRODUCTION

The failure of many vital organs is still one of the leading causes of critical illness and death across the globe. Disease processes affecting vital organs such as the kidneys, liver, heart, and lungs often go undetected as they develop slowly over long periods of time and are accompanied by few, if any, symptoms. Because of this, many patients are only diagnosed when the disease is in its late, and often irreversible, stages, when there are few treatment options and organ transplants become a necessity.

The failure of vital organs is a critical system of illness and death worldwide, and many conditions affecting vital organs are progressive and cause very few symptoms. Because of this, many patients are diagnosed when the conditions are advanced and borderline unmanageable, and organ transplants are the only viable solution. Numerous studies cite organ trans- plant centers as an example of the impact of late recognition, or worse, unrecognized severe organ dysfunction.

Unhealthy living, the environment, chronic diseases, poor genetics, drug abuse and no screening hampers or accelerates organ function. Simply put, organ evaluations lack the early assessment frameworks that filter recycles clinical points. Evans reports the absence of early assessment frameworks on the organ function.

Deciding whether a patient should continue with conventional medical treatment or be evaluated for organ transplant is a complicated clinical process. It requires balancing many clinical parameters, laboratory results, patterns of disease progression, and risk profiles. Chronic kidney disease is an example that can progress to end-stage renal failure [1]–[3]. Liver diseases can progress to cirrhosis or cancer [4]. Heart diseases can progress to end-stage heart failure [6]. And lung diseases can lead to irreversible lung [7]–[10]. The precision with which clinical prognosticators distinguish the phases of a disease is pivotal in designing interventions to extend a patient's life and efficiently allocating an organ transplant to that patient.

For the most part, present clinical techniques depend on a clinician's knowledge and experience, a few relevant diagnostic tests, the clinician's assessment of several images, and one of several established scoring systems. These techniques have been validated clinically, but they are the product of a manual process and do not adequately represent the intricate interplay among many clinical parameters. Most of the existing systems are organ or disease type specific [6]–[10] and do not accommodate the multi organ clinical realities in which they will be applied.

In response to these issues, this research presents a multi-organ machine learning-based decision support framework aimed at the early evaluation of whether treatment is warranted or whether a transplant is needed. This framework uses supervised learning algorithms trained on pertinent clinical and diagnostic data to model the disease severity of each organ. The models will be assessed on their accuracy, precision, recall, and F1 score, ensuring the models will be reliable, interpretable, and scalable.

The intended purpose of the this system is to support clinical staff by giving early, evidence- based assessments of the severity of the disease. This will provide early disease assessments to facilitate timely clinical actions. However, the this system is designed to support clinical staff, and the final determination of treatment will remain with the clinician.

2RELATED WORK

In recent years, there has been a significant increase in the use of machine learning in health care for the prediction and classification of chronic and life-threatening illnesses. As for super- vised learning techniques, several of them—logistic regression, decision tree, random forest, support vector machine, and ensemble methods—have been extensively studied in the context of analyzing clinical datasets and forecasting clinical outcomes. [1, 2, 3, 6, 10].

2.1 Kidney Disease Prediction

Many research studies have focused on machine learning techniques to analyze the laboratory parameters such as serum creatinine, estimated glomerular filtration rate, blood urea and blood electrolyte levels for the purpose of identifying chronic kidney disease [1, 2, 3]. Although these models have shown good performance in diagnosis, the majority of them target disease detection and overlook the evaluation of the severity disease progression and the need for transplant.

2.2 Liver Disease Analysis

Machine learning methods have been widely used for the identification of liver diseases including hepatitis, fibrosis, cirrhosis, and hepatocellular carcinoma. Clinical scoring systems like Model for End-Stage Liver Disease (MELD), Milan criteria and staging systems are routinely used in practice [4, 5]. However, these methods are typically used in an isolated fashion and are poorly integrated with predictive machine learning models.

2.3 Heart Disease Prediction

Predicting heart disease has been a leading focus of research within the field of medical ma- chine learning. Some of the most common features used by the models include age, blood pressure, cholesterol, findings from electrocardiograms, and responses from exercise tests [6, 7]. These models adequately identify risk of cardiac events; however, they almost never capture the progression of disease to transplant-level severity.

2.4 Lung Disease Assessment

Pulmonary function tests, clinical signs, image-based analytics, and novel electronic nose technologies for breath analysis have been extensively researched for the application of machine learning in lung disease detection [8, 9, 10, 11, 12, 13, 14]. While these studies contribute to the early identification of lung diseases, relatively little research has been done on assessing the need for transplants, or the incorporation of these studies in clinician decision support systems.

2.5 Limitations of Existing Work

Although there have been improvements in different organ systems, other studies have shown some limitations such as: focusing on predicting in just one organ, not fully addressing the role of treatment vs transplant decision-making, overly simplified binary classification, and lack of clinical workflow integration. These limitations highlight the importance of providing a comprehensive multi organ integrated clinical decision support system that enables early and interpretable severity assessment.

3 MATERIALS

This study uses multiple clinically validated datasets corresponding to different organ systems to construct a unified, multi-organ decision-support framework. [1, 2, 6, 10].

3.1 Heart Disease Dataset [26]

This dataset (Heart.csv) consists of 70, 000 patient records collected from clinical examinations, where the target variable represents binary cardiovascular disease status (Class 0: Absent, Class 1: Present). The dataset exhibits balanced class distribution with 35, 021 negative cases (50.03%) and 34, 979 positive cases (49.97%) [6, 7]. Each record contains twelve clinical attributes including age (years), gender (1=Female, 2=Male), height (cm), weight (kg), systolic blood pressure (mmHg), diastolic blood pressure (mmHg), cholesterol level (1=Normal, 2=Above Normal, 3=Well Above Normal), glucose level (1=Normal, 2=Above Normal, 3=Well Above Normal), smoking status, alcohol intake, and physical activity level. The dataset exhibits balanced class representation across both positive and negative cases, providing a robust evaluation environment without severe class imbalance problems frequently encountered in medical datasets.

3.2 Chronic Kidney Disease Dataset [21]

Chronic Kidney Disease (CKD) patient records, along with CKD's structured clinical dataset containing demographic information, clinical laboratory data, and physical measures (i.e., data) relevant to the renal

assessment, comprise the dataset utilized for kidney disease analysis [1, 2, 3]. A multi-class classification analysis focused on the prediction of disease progression was enabled by the datasets, which are defined by discrete stages of CKD. For the sake of obtaining an unbiased result, the CKD data distribution (i.e. class) was maintained when the data was allocated to training and testing divisions. Obtained from the dataset, the collection of renal dysfunction data, having supports and data reflective of all levels of detail, all adds to and includes of (i.e. data) the early stage loss of function and the later stage complete kidney failure.

3.3 Liver Disease Datasets

Two independent datasets were employed to model liver-related disorders at different clinical stages, covering both general liver disease screening and advanced malignancy outcomes [4, 5].

3.3.1 Indian Liver Patient Dataset (ILPD) [24]

The Indian Liver Patient Dataset (ILPD) was used for binary liver disease classification. The dataset contains 117 patient records described by 11 routinely measured clinical attributes, including Age, Gender, Total Bilirubin, Direct Bilirubin, Alkaline Phosphatase (ALP), Alanine Aminotransferase (ALT), Aspartate Aminotransferase (AST), Total Proteins, Albumin, Albumin–Globulin Ratio, and a binary outcome label indicating liver disease presence.

The class distribution is moderately imbalanced, with 87 healthy individuals and 30 patients diagnosed with liver disease.

3.3.2 Hepatocellular Carcinoma Dataset [25]

The Hepatocellular Carcinoma dataset represents a real-world clinical cohort collected at the Hospital and University Centre of Coimbra, Portugal. It includes 165 patient records characterized by demographic, laboratory, tumor-related, lifestyle, and comorbidity attributes selected according to established clinical guidelines [4]. The prediction target corresponds to one year survival outcome encoded as 0 (death) and 1 (survival). The dataset exhibits moderate class imbalance with 63 non-survivors and 102 survivors.

3.4 Lung Disease (COPD) Datasets

The analysis of Chronic Obstructive Pulmonary Disease (COPD) has been supported by two complementary datasets that allow both early diagnosis and evaluation of the severity of the disease [8, 10, 11, 12, 13, 14].

3.4.1 Stage-1 Respiratory Audio Dataset [22]

The first dataset consists of recordings of respiratory sounds and includes participants who are confirmed to have COPD, smokers, and healthy control subjects. The raw respiratory signals were processed to obtain the numerical feature elements of the signals, mean, variance, signal minimum and maximum, median, RMS, signal energy, skewness, and kurtosis. These features reflect the variability of airflow and describe the abnormal breathing cycles that are characteristic of obstructive pulmonary disorders. The samples were labeled as either COPD, smoker, or healthy control.

3.4.2 Stage-2 Clinical COPD Dataset [23]

The second dataset consists of the organized clinical records of 230 COPD patients. The records captured the demographic, lifestyle, and physiological attributes, including age, sex, BMI, smoking history, occupational exposure, presence of dyspnea, respiratory rate, level of oxygen saturation, characteristics of sputum, and comorbid conditions. The severity of the disease was determined according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification, which ranges from GOL[14].

4 METHODS

We present here the collaborative methodological approach for the multi-organ disease pre- diction involving the heart, kidneys, liver, and lungs [1, 2, 6, 10].

4.1 Data Preprocessing and Feature Preparation

All datasets underwent standardized preprocessing to enhance data quality and model robustness. Class imbalance was carefully evaluated. Although synthetic oversampling (SMOTE) was explored, it was excluded from the final pipeline due to increased variance and reduced generalization performance, particularly in ensemble models.

4.2 Machine Learning Model Development

A diverse set of supervised learning algorithms was evaluated to capture both linear and non-linear relationships in the clinical data. These included Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbors, AdaBoost, Gradient Boosting, Naïve Bayes, and XGBoost. Hyperparameters were tuned using consistent empirical strategies across all datasets to maintain stability and avoid dataset-specific overfitting.

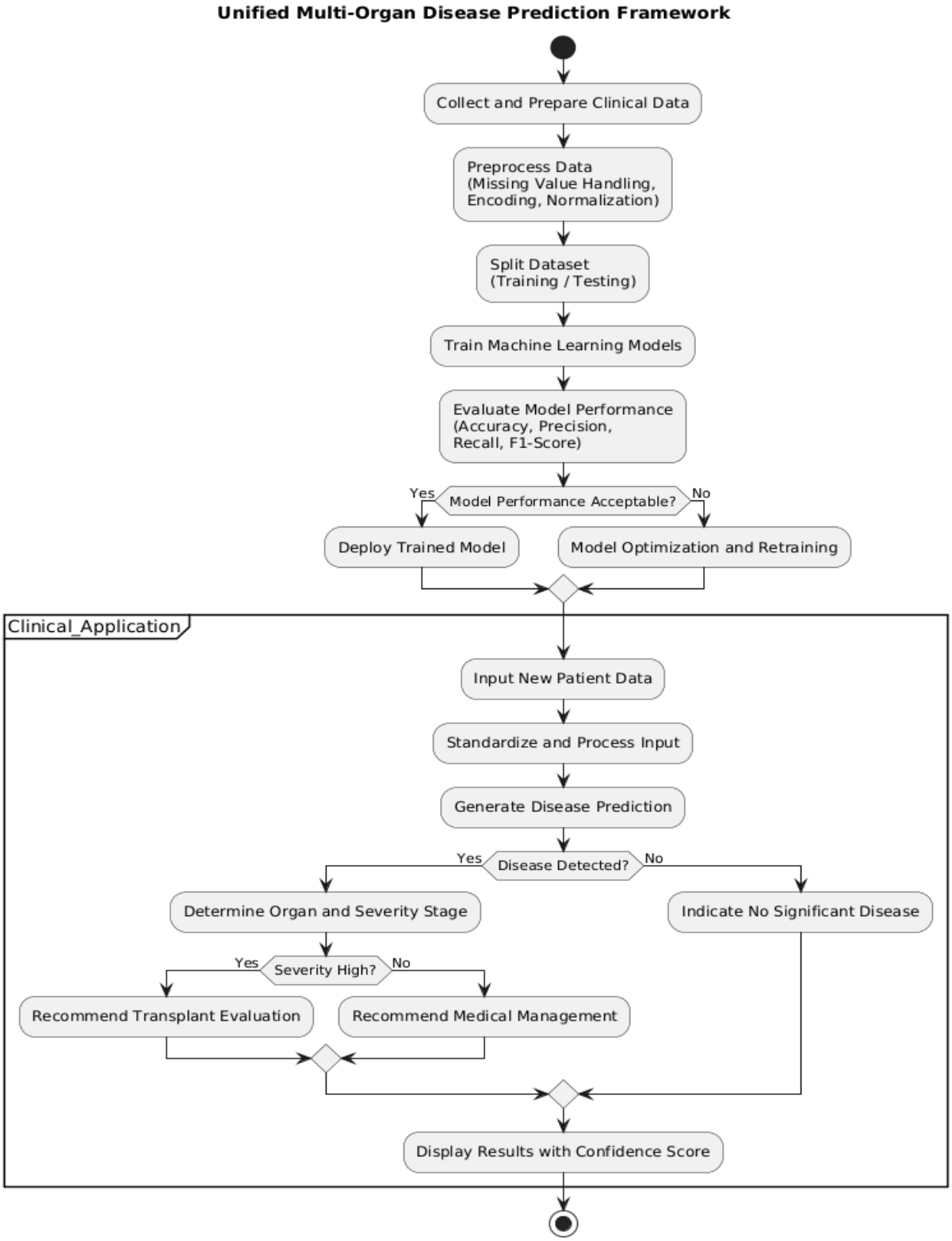

Fig 1 : Activity Diagram

4.3 Ensemble Learning Strategies

Incorporating several ensemble learning methods to improve predictive stability and minimize model bias is a common strategy. Individual base classifiers were combined to create a hard voting and a soft voting ensemble. As with previous studies in healthcare ensemble modeling, [2, 12], probability stacking-based ensembles were used for datasets requiring multiclass severity prediction. The stacking ensemble used a heterogeneous base learner with a meta-classifier designed to output final predictions based on the learner's outputs. This method takes advantage of the complementary decision boundaries of the individual models to reduce variance and improve stage-wise generalization across the severity classes of the disease.

4.4 Organ-Specific Modeling Adaptations

While the core pipeline remained consistent, minor adaptations were applied to address organ- specific characteristics:

- **Heart Disease**: A multi-class severity prediction framework was implemented, with emphasis on macro-averaged metrics to account for severe class imbalance among advanced disease stages [6, 7].
- **Chronic Kidney Disease**: Multi-class CKD stage prediction was performed, prioritizing macro recall to minimize under-detection of advanced renal impairment, which is clinically critical [1, 3].
- **Liver Disease**: Binary classification was applied for the ILPD dataset, while survival out- come prediction was performed for the HCC dataset, guided by clinical transplant allocation considerations [4, 5].
- **COPD**: A sequential two-stage architecture was developed. Stage-1 performed respiratory sound-based screening, followed by Stage-2 clinical severity classification based on GOLD criteria [8, 10, 11, 12, 14].

4.5 Model Evaluation Metrics

Model performance was evaluated using Accuracy, Precision, Recall, and F1-score, which are widely adopted evaluation metrics in medical machine learning studies [1, 2, 6]. For imbalanced and multi-class datasets, macro-averaged recall and macro-averaged F1-score were prioritized to ensure equitable assessment across minority classes.

5 RESULTS

This section presents the experimental results obtained for each organ-specific prediction task using the unified machine learning pipeline. Performance is reported separately for cardio- vascular, renal, hepatic, and respiratory datasets to preserve clinical clarity and task-specific interpretation.

5.1 Cardiovascular Disease (Heart) Classification Performance

The heart disease prediction experiments were carried out using the Heart Patient Dataset avail- able on Kaggle [26]. The dataset includes cardiovascular-related clinical features and labeled outcomes for binary classifications of the heart disease risk. Table 1 summarizes the performance of all evaluated models on the cardiovascular disease dataset (n=14, 000 test samples).

Table 1: Cardiovascular Disease (Heart) Binary Classification Performance

Model	**Accuracy**	**Macro Recall**	**Macro F1-Score**
XGBoost	**73.36**	**0.73**	**0.73**
AdaBoost	72.23	0.72	0.72
Decision Tree	71.78	0.72	0.72
Logistic Regression	71.36	0.71	0.71
Random Forest	71.15	0.71	0.71
K-Nearest Neighbors	64.83	0.65	0.65
Naive Bayes	58.91	0.59	0.55
Support Vector Machine	49.73	0.50	0.35
Gradient Boosting	73.07	0.73	0.73
Stacking Ensemble	73.29	0.73	0.73
Hard Voting Ensemble	73.14	0.73	0.73
Soft Voting Ensemble	72.78	0.73	0.73

5.2 Chronic Kidney Disease (CKD) Stage Prediction Results

The experiments were conducted using the publicly available CKD dataset with stage annotations from Kaggle [21]. Table 2 presents the performance comparison of individual and ensemble models for multi-class CKD stage prediction.

Ensemble-based models consistently outperformed individual classifiers. The stacking ensemble achieved the highest overall performance, demonstrating balanced sensitivity across early and advanced CKD stages.

Table 2: Performance Comparison of Models for CKD(Kidney) Stage Classification

Model	**Accuracy**	**Macro Recall**	**Macro F1**
Logistic Regression	0.934	0.892	0.910
Decision Tree	0.991	0.964	0.972
Support Vector Machine	0.895	0.831	0.851
K-Nearest Neighbors	0.470	0.439	0.452
Random Forest	0.974	0.884	0.901
Gradient Boosting	0.993	0.969	0.978
AdaBoost	0.968	0.830	0.820
Naïve Bayes	0.718	0.759	0.690
XGBoost	0.985	0.964	0.973
Soft Voting Ensemble	0.991	0.975	0.980
Hard Voting Ensemble	0.989	0.961	0.970
Stacking Ensemble	**0.994**	**0.977**	**0.982**

5.3 Liver Disease Prediction Results

5.3.1 ILPD Liver Disease Classification

Experiments were carried out using the Indian Liver Patient Dataset (ILPD), which we sourced from the UCI Machine Learning Repository [249]. The dataset includes clinical and biochemical features of both liver patients and healthy subjects, which supports binary classification for the task of liver disease detection. Data preprocessing and normalization techniques have been carried out before model training. The performance of different classifiers on the ILPD dataset is summarized in Table 3.

Table 3: ILPD Liver Disease Classification Performance

Model	**Accuracy**	**Macro Recall**	**Macro F1**
Logistic Regression	0.66	0.76	0.65
Decision Tree	0.71	0.67	0.65
Random Forest	0.68	0.74	0.66
Support Vector Machine	0.61	0.70	0.60
AdaBoost	0.75	0.70	0.69
Gradient Boosting	0.73	0.60	0.60
XGBoost	0.77	0.63	0.64
K-Nearest Neighbors	0.74	0.67	0.66
Gaussian Naïve Bayes	0.51	0.67	0.51
Hard Voting	**0.82**	**0.82**	**0.79**
Soft Voting	0.79	0.79	0.79
Stacking	0.76	0.76	0.76

The Hard Voting Ensemble achieved the highest accuracy and balanced recall, making it the most reliable model for ILPD-based liver disease screening.

5.3.2 Hepatocellular Carcinoma (HCC) Liver Survival Prediction

The experiments on survival prediction for HCC were done with the dataset from the UCI Ma- chine Learning Repository [25] on Hepatocellular Carcinoma *(HCC)*. The dataset includes the clinical and lab data of patients with *HCC* and allows us to perform a binary classification for the prediction of the survival outcomes. Before training the models, the data was preprocessed and normalized. Table 4 presents the performance of classifiers on the HCC(Liver) dataset.

Table 4: HCC(Liver) Survival Prediction Performance

Model	**Accuracy**	**Macro Recall**	**Macro F1**
Logistic Regression	0.79	0.80	0.78
Decision Tree	0.82	0.82	0.81
Random Forest	0.91	0.90	0.90
Support Vector Classifier	0.82	0.82	0.81
K-Nearest Neighbors	0.70	0.68	0.68
AdaBoost	**0.97**	**0.96**	**0.97**
Gradient Boosting	0.91	0.90	0.90
Hard Voting	0.94	0.94	0.94
Soft Voting	0.64	0.64	0.64
Stacking	0.70	0.70	0.70

AdaBoost achieved the highest individual accuracy, while the voting ensemble provided stable and well-balanced predictions across survival classes.

5.4 COPD Prediction Results

5.4.1 Stage-1 COPD(Lungs) Screening Performance

The experiments for Stage-1 COPD screening were carried out based on an openly accessible dataset COPD GOLD containing respiratory sounds of clinically diagnosed patients [22]. Since the dataset contains audio samples of varying COPD severity levels, it is possible to use supervised classification with the acoustic features extracted from the samples. For each audio samples, respiratory signal was processed, and features that are descriptive of the varying characteristics of the breath sounds in the time and frequency domains were extracted and used. This dataset is a good benchmark for testing of all the machine learning and ensemble-based approaches for early screening of COPD. Stage-1 screening performance using respiratory sound features is summarized in Table 5. Tree-based ensemble models clearly outperform linear and probabilistic approaches. The Extra Trees classifier achieves the highest accuracy and F1-score, indicating superior capability in modeling non-linear breath signal characteristics.

Table 5: Stage-1 COPD(Lungs) Screening Performance

Model	**Accuracy**	**Recall**	**F1-score**
Logistic Regression	0.8289	0.8289	0.8264
Linear SVM	0.8190	0.8190	0.8151
RBF SVM	0.8157	0.8157	0.8126
Decision Tree	0.8421	0.8421	0.8420
Random Forest	0.8880	0.8880	0.8878
Extra Trees	**0.9029**	**0.9029**	**0.9028**
Decision Tree	0.8815	0.8815	0.8816
Gradient Boosting	0.8782	0.8782	0.8784
AdaBoost	0.7878	0.7878	0.7829
Gaussian Naïve Bayes	0.7153	0.7153	0.7103
K-Nearest Neighbors	0.8323	0.8323	0.8292
Voting (Hard)	0.8864	0.8864	0.8860
Voting (Soft)	0.8749	0.8749	0.8746
Stacking	0.8717	0.8717	0.8709

5.4.2 Stage-2 COPD(Lungs) Severity Classification

The experiments in classifying severity of Stage-2 COPD were carried out with the COPD GOLD dataset [23]. In this phase, we engage with the multi-class classification of the different levels of GOLD severity, making it more difficult as there are overlapping audio patterns among the stages. This difficultly is why we see lower performance in comparison to the Stage-1 screenings. Table 6 summarizes the performance of models used for COPD severity prediction based on GOLD stages.

Table 6: Stage-2 COPD(Lungs) Severity Classification Performance

Model	Accuracy	Weighted Recall	Weighted F1
Logistic Regression	0.5802	0.5802	0.5719
Linear SVM	0.5535	0.5535	0.5368
RBF SVM	0.5849	0.5849	0.5758
Decision Tree	0.5493	0.5493	0.5450
Random Forest	0.6295	0.6295	0.6208
Extra Trees	**0.6386**	**0.6386**	**0.6344**
Decision Tree	0.6161	0.6161	0.6128
Gradient Boosting	0.6158	0.6158	0.6081
AdaBoost	0.4728	0.4728	0.4539
Gaussian Naïve Bayes	0.5668	0.5668	0.5522
K-Nearest Neighbors	0.5489	0.5489	0.5315
Voting (Hard)	0.6072	0.6072	0.6003
Voting (Soft)	0.6338	0.6338	0.6266
Stacking	0.5711	0.5711	0.5572

6 FEATURE IMPORTANCE ANALYSIS

6.1 Heart Disease

For heart dataset Age proves to have the greatest positive correlation to cardiovascular disease, comments on the positive correlation which is noticed by experts as the greatest predominance. The level of cholesterol and the body weight shows strong positive correlations which suggests the contribution of these variables toward the levels of CVD risk. When an expert analyzes blood glucose levels, they observe a positive correlation to disease, but the correlation is not strong, which indicates a positive correlation and suggests a strong relevance. The clinical variables of systolic and diastolic blood pressures both have clinical positive correlations and reiterate the well documented relation between hypertension and cardiovascular disease. Demographic variables like, sex, and lifestyle factors like, alcohol consumption, smoking, and stature, show weak positive correlations which suggest and infer a lower direct contribution and impact in this dataset. Notably, physical activity shows a negative correlation with cardiovascular disease, emphasizing its protective effect against cardiac risk. Overall, the correlation analysis highlights that biological and metabolic parameters dominate cardiovascular risk assessment, while lifestyle factors contribute indirectly. These findings support the clinical validity and interpretability of the risk analysis framework used in this decision support system.

6.2 Chronic Kidney Disease (CKD) Stage

Feature importance was evaluated using the Random Forest classifier to examine the contribution of each variable to CKD stage prediction. The most significant predictor for the model is the Glomerular Filtration Rate (GFR). This is also clinically relevant as chronic kidney disease (CKD) staging is determined with respect to ranges of GFR levels. Other important clinical metrics for the classification include serum creatinine, blood pressure, BUN, pH and calcium levels of the urine. In the lifestyle factors, smoking and drinking alcohol, diet and physical activity have much less relevance.

6.3 Liver Disease

Biomarkers that are liver specific and demonstrate the highest strength and the most clinical relevance in relation to the liver disease risk are the direct bilirubin and total bilirubin levels, ALT, AST, and alkaline phosphatase. Indicators of the proteins such as albumin, and the albumin to globulin ratio, are all positive indicating that liver synthetic function is impaired. Age and sex/influence are however moderates.

6.4 Chronic Obstructive Pulmonary Disease (COPD) – Stage 1: Respiratory Audio Screening

The importance values indicate the relative contribution of each feature toward classification of COPD, smokers, and healthy controls. The minimum amplitude (Min) exhibits the highest contribution, suggesting that extreme airflow variations play a critical role in distinguishing obstructive patterns. Higher-order statistical measures such as skewness and kurtosis also demonstrate strong influence, indicating that distribution asymmetry and peak sharpness are important indicators of irregular breathing sounds. Furthermore, the features related to variability, such as standard deviation (Std), RMS, and energy, show some importance, as they may differ based on signal intensity and turbulence of the airflow across the different subject groups.

6.4.1 Stage-2 Clinical COPD Dataset

Using the Extra Trees classifier, feature importance was determined to ascertain which features were the best predictors for the severity of COPD. the dyspnea mMRC showed to be the most important feature due to its great correlation with GOLD staging. Clinical relevance of the mMRC dyspnea scoring system, respiration and FEV1 significantly contributes to the limitation of airflow. While other clinical and demographic factors like age, crucialness of the mMRC dyspnea scoring system, respiration, oxygen saturation and FEV1, and the history of smoking and exposure to occupational hazards, have been and continue to be relatively important, the co-morbidities have shown to be significantly less important.

7 DISCUSSION

We used several datasets for implementing this model. It is observed that for Heart dieses BMI and cholesterol features plays significant role. For Kidney dieses GFR feature plays significant role. For Liver dieses Bilirubin feature plays significant role. For Lung dieses Min and mMRC feature plays significant role.

8 Limitations and Future Work

For the future of this technology to be realized, there needs to be real-time integration with other systems i.e. hospital information systems, electronic health record systems and diagnostic platforms. This will enable the model to be continuously updated and the provision of real-time clinical decision support during the provision of care will greatly enhance the effectiveness of this model within the real-life clinical setting.

9 Conclusion

From Table 1 to 6 it is clear that for Heart disease XGboost gives better performance, for Kidney disease Stacking Ensemble technique gives better performance, for Liver disease Hard voting Ensemble technique gives better performance, for Lungs disease Extra Trees gives better performance. The system's highly scalable architecture allows future hospital information system integration as well as the national transplant registry. The this system is useful and advanced for modern environments of transplant and disease management. The this system is designed to improve the digital healthcare. It helps to improve medical decision-making, especially for organ transplantation. It is also designed with the flexibility to be integrated into hospital information systems in the future, as well as national transplant registries. Responsive to the advanced nature of digital healthcare, the system is a necessity in modern environments for effective management of transplant.

References

[1] S. Akter, A. Habib, M. A. Islam, M. S. Hossen, W. A. Fahim, and P. R. Sarkar, "Comprehensive performance assessment of deep learning models in early prediction and risk identifica- tion of chronic kidney disease, " in *Proc. 5th Int. Conf. Electrical Engineering and Informa- tion & Communication Technology (ICEEICT)*, 2021, pp. 1–6.

[2] M. S. Haque, M. S. Amin, S. Ahmad, M. A. Sayed, A. Raihan, and M. A. Hossain, "Predicting kidney failure using an ensemble machine learning model: A comparative study, " in *Proc. 10th Int. Conf. Electrical Engineering, Computer Science and Informatics (EECSI)*, 2023, pp. 31–37.

[3] M. K. Grewal, B. Mohapatra, A. Kumar, G. Mehta, M. Saxena, and P. K. Sarangi, "A machine learning approach to early detection and classification of kidney disease, " in *Proc. 3rd Int. Conf. Communication, Security, and Artificial Intelligence (ICCSAI)*, 2025.

[4] M. S. Reddy *et al.*, "National liver allocation policy—Consensus document by the Liver Transplantation Society of India for a nationally uniform system of allocation of deceased donor liver grafts, " *J. Clin. Exp. Hepatol.*, vol. 13, no. 2, pp. 303–318, Mar.–Apr. 2023, doi: 10.1016/j.jceh.2022.12.001.

[5] United Network for Organ Sharing, "Updates to medical urgency scoring of liver transplant candidates in effect, " Jul. 13, 2023. [Online]. Available: https://unos.org/news/policy- changes/updates-to-medical-urgency-scoring-of-liver-transplant-candidates-in-effect/

[6] T. Nagaraj and B. Bhaskara Rao, "An effective heart disease detection and severity level classification model using machine learning and hyperparameter optimization methods, " in *Proc. 6th Int. Conf. Computing Methodologies and Communication (ICCMC)*, 2022.

[7] F. M. A. Hossain *et al.*, "A multi-task neural network architecture for renal dysfunction pre- diction in heart failure patients with electronic health records, " *IEEE J. Biomed. Health Inform.*, vol. 24, no. 8, pp. 2303–2310, 2020.

[8] S. D. Almeida, A. C. Silva, A. C. Paiva, and M. Gattass, "Prediction of disease severity in chronic obstructive pulmonary disease using quantitative analysis of chest computed to- mography, " *Comput. Biol.*

Med., vol. 44, pp. 117–126, 2014.
[9] M.-C. Elie-Turenne *et al.*, "Lung injury prediction score for the emergency department, "
Crit. Care Med., vol. 40, no. 5, pp. 1480–1486, May 2012.
[10] M. H. Al-Sheikh, O. Al Dandan, A. S. Al-Shamayleh, H. A. Jalab, and R. W. Ibrahim, "Multi- class deep learning architecture for classifying lung diseases from chest X-ray and CT im- ages, " *Scientific Reports*, vol. 13, Art. no. 19373, 2023.
[11] Y. Li, X. Wei, Y. Zhou, J. Wang, and R. You, "Research progress of electronic nose technology in exhaled breath disease analysis, " *Microsyst. Nanoeng.*, vol. 9, Art. no. 129, Oct. 2023, doi: 10.1038/s41378-023-00594-0.
[12] V. A. Binson, M. Subramoniam, and L. Mathew, "Detection of COPD and lung cancer with electronic nose using ensemble learning methods, " *Clin. Chim. Acta*, vol. 523, pp. 231– 238, Dec. 2021, doi: 10.1016/j.cca.2021.10.005.
[13] C. Cuastumal Vásquez, C. Durán Acevedo, and J. Rueda Paredes, "Electronic nose dataset for COPD detection from smokers and healthy people through exhaled breath analysis, " *Data in Brief*, vol. 38, Art. no. 106767, Jun. 2021, doi: 10.1016/j.dib.2021.106767.
[14] A. Shojaei and M. Delrobaei, "Severity classification of chronic obstructive pulmonary disease in intensive care units: A semi-supervised approach using MIMIC-III dataset, " arXiv:2504.18593, 2025.
[15] R. P. Bagawade, R. Manda, J. Gavit, V. Karanjakar, and A. Lohar, "Heart Disease Prediction Using Deep Learning, " *International Research Journal of Modernization in Engineering Technology and Science (IRJMETS)*, vol. 6, no. 5, pp. 8186–8191, May 2024.
[16] R. P. Bagawade and D. R. Thirupurasundari, "A Review Data Pre-Processing Techniques in Machine Learning, " in *Proc. Hinweis Second International Conference*, RTET, Grenze International Journal of Engineering and Technology, vol. 48, pp. 3937–3946, Dec. 2024.
[17] R. P. Bagawade and D. R. Thirupurasundari, "Disease Prediction Using Machine Learning, " *Journal of Engineering and Technology Management (JETM)*, vol. 74, no. 1, pp. 1–9, Dec. 2024.
[18] R. P. Bagawade and D. R. Thirupurasundari, "A Review of Feature Selection Techniques for Machine Learning Algorithms used for Disease Prediction, " in *Proc. Fifth International Conference on Recent Trends in Science, Engineering and Technology (ICRTSET)*, ISTE, vol. 48, Special Issue 2, pp. 7–13, Nov. 2025.
[19] R. Bagawade, S. Pehere, S. Gadekar, and U. Kulkarni, "Liver, Kidney, Heart and Diabetes Disease Prediction using Machine Learning, " *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 13, no. 3, pp. 2728–2734, Mar. 2025.
[20] R. P. Bagawade and D. R. Thirupurasundari, "Ensemble Machine Learning Techniques, " *International Journal of Emerging Technology and Advanced Engineering (IJETAE)*, vol. 15, no. 4, pp. 15–23, Apr. 2025.
[21] A. Hirlekar, "CKD (Chronic Kidney Disease) Dataset with Stages, " Dataset on Kag- gle, 2025. Available: https://www.kaggle.com/datasets/aryanhirlekar/ckdchronic-kidney- disease-dataset-with-stages (accessed Feb. 12, 2026).
[22] C. M. Durán Acevedo, C. A. Cuastumal Vasquez, and J. K. Carrillo Gómez, "Electronic nose dataset for COPD detection from smokers and healthy people through exhaled breath analysis, " *Data in Brief*, vol. 35, p. 106767, 2021, doi:10.1016/j.dib.2021.106767.
[23] sawdi777, "COPD GOLD 230 Patients, " Kaggle Dataset, 2025. Available: https://www.kaggle.com/datasets/sawdi777/copd-gold-230-patients (accessed Feb. 12, 2026). :contentReference[oaicite:0]index=0
[24] D. Dua and C. Graff, "UCI Machine Learning Repository, " University of Califor- nia, Irvine, School of Information and Computer Sciences, 2019. [Online]. Available: https://archive.ics.uci.edu (accessed Feb. 12, 2026).
[25] D. Dua and C. Graff, "UCI Machine Learning Repository, " University of Califor- nia, Irvine, School of Information and Computer Sciences, 2019. [Online]. Available: https://archive.ics.uci.edu (accessed Feb. 12, 2026).
[26] O. Midon, "Heart Patient Dataset, " Kaggle Dataset, 2025. Available: https://www.kaggle.com/datasets/omidon/heart-patient-dataset (accessed Feb. 12, 2026).