A Hybrid Machine Learning Framework with Optimized Feature Selection for Augmenting Survival Prediction via Synthetic Match Generator

Padmini Kuppala[1], Dr. Nilesh V. Ingale[2], Dr. V. Aruna[3]

[1,3]Department of CSE, Sreenidhi Institute of Science and Technology, Ghatkesar, Hyderabad

[2]Department of CSE, Vikrant University Gwalior-India

DOI: https://doi.org/10.47772/IJRISS.2026.1026EDU0131

Received: 18 February 2026; Accepted: 24 February 2026; Published: 18 March 2026

ABSTRACT

Heart surgery is the most important thing that can be done for kids with end-stage heart failure, but there is still a big problem with death one year after the transplant. It is very important to get this mortality risk right in order to match donors and recipients more effectively and improve patient results. In this work, we use the ICU heart transplant expiration dataset to estimate the risk of death in pediatric heart transplant patients after one year. We suggest a new method that uses advanced feature selection and group methods to make predictions more accurate. Using Chi-squared tests to pick out key traits and combining multiple classifiers for accurate predictions are part of the method. The results show that the suggested Voting Classifier, which uses both Boosted Decision Tree and ExtraTree models, works very well, as it gets 100% of the votes right. This method is a quick and accurate way to guess the chance of death within a year. It gives doctors useful information for better patient care and finding the best match between recipient and donor in pediatric heart transplants.

"Index Terms - Machine learning algorithms, deep learning, classification, sleep disorder, Voting algorithm".

INTRODUCTION

"Heart transplantation (HTx)" is now a technique that can save the lives of children with advanced heart failure. Even though they only make up about 10% of all heart transplants done each year, the number of pediatric HTx cases has slowly grown over the last few decades. In 2020 alone, more than 450 of these surgeries will be done in the United States. This growth is due to improvements in medical technology and surgical methods. However, there are still problems, especially when it comes to lowering the death rate one year after transplantation, which is still very high [7]. There aren't many good organs available for pediatric HTx, which makes it even more difficult. This adds to the serious problem of people dying while they are on the waiting list. A lot of pediatric heart donors are thrown away because it's hard to tell if the organs are good enough and if the recipients will be compatible [7]. This shows the need for better donation utilization strategies.

To get better results in pediatric HTx, people have been looking into what makes a transplant work and making tools for visualizing data better to help doctors make decisions. Even with all of these attempts, matching donors and recipients is still a very subjective process that depends on a lot of different factors about both parties, such as medical, physiological, and demographic factors [4, 10]. So, making reliable prediction models to look at what will happen after a transplant is necessary to improve systems for allocating organs and help doctors make better decisions [19].

Prediction models have been very helpful in making decisions about heart transplants. One measure that is often used for allocation is the "Heart Transplantation Survival Score (HTSS)", which was made by the "United Network for Organ Sharing (UNOS)" in the US. The HTSS takes into account things like the person's age, illness, level of function, and any other health problems they may have, like diabetes or kidney disease. It also looks at things about the donor, like their age, the reason of their death, and whether their blood type is compatible with the recipient's. This score will provide a numerical value of the probability of surviving following transplant and this will enable donation centers to determine which of the patients awaiting transplants should be assisted initially [13].

The "Eurotransplant International Foundation developed the Eurotransplant Donor Risk Index (ET-DRI)", which also examines aspects related to both the recipient and the donor so as to allow individuals to make decisions regarding a transplant. The models such as HTSS and ET-DRI demonstrate that predictive analytics are essential in the field of heart surgery. These models are not flawless, however, as they may not consider all factors, which influence the effectiveness of transplant. Better prediction accuracy based on current data-driven approaches is crucial to enhancing survival rates of recipients, optimizing organ allocation, and resolving the issues that continue to arise in pediatric heart transplantation [2].

ML will enable one to reliably determine the characteristics of donors in pediatric heart transplants, thereby rendering it far simpler to see how the transplant will proceed. This approach addresses issues in donating organs because it makes the decisions more effective and increases the survival rates of patients [20].

Related Work

Heart transplantation has now emerged as a life saving procedure among children with end-stage heart failure, yet issues such as transplantation allocation and post-transplantation death remain. The use of ML as a way to address these issues has become a fascinating topic since it can allow doctors to make decisions and promote better results through the provision of predictive models.

Ashfaq et al. [1] applied the data on the "United Network of Organ Sharing (UNOS)" database and used the ML model to predict the probability of death one year after a child heart transplant. They found that ML models could be superior at outcome prediction compared with traditional statistical approaches because they consider many donor and recipient characteristics. Miller et al. [15] applied the ML techniques, including the Random Forests and Neural Networks, to enhance death prediction in pediatric heart transplantation. They demonstrated that the ML methods were capable of providing more accurate predictions than conventional scoring systems. This can imply replacing the usage of traditional models to the usage of data-driven models to assist physicians in decision-making.

Killian et al. [12] examined national registry data to make an estimate of what would occur to children undergoing heart transplants. Using their research, they demonstrated the significance of preprocessing data, selecting the appropriate features, and optimizing the model to achieve the required accuracy of predictions when comparing various methods of ML. In the study, it was also noted that ML models are adaptable to new data and hence, suitable in changing clinical situations. This is elaborated by Gotlieb et al. [9], who discuss the application of ML in solid organ transplants including heart transplantation. They examined the potential of ML models to assist in patient selection, organ- organ matching, and post-surgery care. They identified certain methods how ML could reduce the variation of clinical decisions and enhance outcomes.

Chebli et al. [6] considered the application of the semi-ML in medicine, and the way that it is applicable in heart transplants. The issue of an insufficient amount of labeled data in healthcare was resolved with their work because semi-supervised learning was used to teach prediction models well. This technique is particularly effective in the case of pediatric heart transplants, where data is frequently hard to access. When models can identify significant trends in both labeled and unlabeled data, then it is preferable that they be robust and generalizable.

Miller et al. [16] investigated the temporal dynamics of the accuracy of ML models in predicting the outcome of heart transplant. They claimed that the prediction of the future becomes untrue as time passes by due to changes in the clinical practices, patient population, and organ supply in the ML models. Their analysis had proposed that the models must be continuously re-trained on new data to ensure that performance remains high. This finding can be of great significance to the pediatric heart transplants because a quick advancement in medical treatment and alteration in the manner of giving out the organs may render older models obsolete.

SHAP (SHapley Additive Explanations), the generation of Lundberg and Lee [14] is a coherent means of interpreting statements produced by ML models. Such an approach is particularly effective in a healthcare context such as children heart transplants because it is essential that physicians know how various characteristics influence the final results to make optimal choices. SHAP guides physicians to interpret the workings of ML predictions by providing them with uniform and intuitive values of the feature importance. This creates honesty and receptivity. The technique is very common where interpretability is of the essence. It eliminates one of the primary issues that render the successful application of complex ML models to clinical practice complicated.

A systematic study by Naruka et al. [17] demonstrated the application of ML and AI found in heart transplants. Their study examined the various processes of ML, such as supervised and unsupervised learning, and their usefulness in predicting the outcome of transplants, discovering the optimal match between donors and recipients and assessing the quality of organs. The authors emphasized the idea that the ML can solve the issues with the existing methods of allocation and improve the situation in the long-term perspective. Nevertheless, they also highlighted such issues as data standardization and ethical issues. To harness the potential of ML to its full extent in cardiac transplantation, they urged doctors and data scientists to collaborate.

Yang et al. [18] conducted a critical review of deep semi-supervised learning, specifically, how it can be applied in cases where there are limited amounts of labelling data, such as that of infant heart transplantation. They discussed newer approaches, including generative models, consistency regularization, which improve learning with labeled and unlabeled data. This approach is particularly effective in clinical environments where tagged data is difficult to locate, and is also very expensive to retrieve. Deep semi-supervised learning is capable of enhancing model generalization and performance with good use of unlabeled data. This renders it an excellent approach to enhance ML applications in medical care.

MATERIALS AND METHODS

We propose a more sophisticated method, in which a new combination of ML and feature selection techniques is used to estimate the death of a child a year after heart transplantation. The system will use the ICU heart transplant expiration data to identify significant variables that increase the risk of post-donation death. We will use the “Chi-squared (Chi2)” method to filter out irrelevant traits and make sure the model focuses on the most important predictors. This will improve the accuracy of our predictions. Then, to make predictions, we will use several ML methods, “such as K-Nearest Neighbors (KNN) [11], Logistic Regression [8], and Random Forest [3]”. We will also look into ensemble methods, especially the Voting Classifier, which uses the best features of both Boosted Decision Trees and ExtraTree models to make the system work better by combining their strengths. A big part of the suggested system will be using semi-supervised learning methods. In these methods, fake cases are made by matching donors and recipients in a way that looks a lot like real cases. This will let the system use data that hasn't been labeled, which will make it better at making accurate guesses. The suggested method is meant to improve clinical outcomes and make it easier for donors and recipients to find each other.

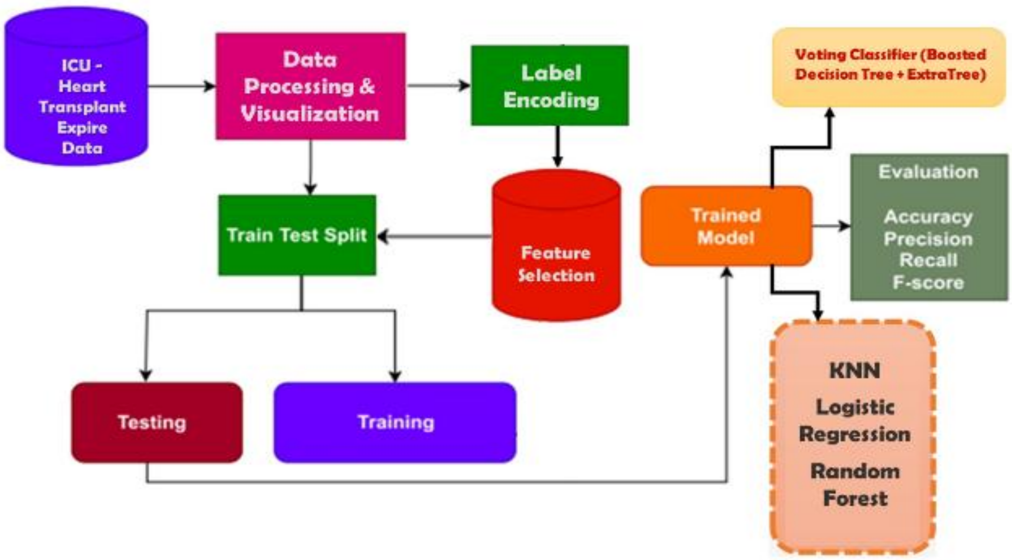


Fig.1 Proposed Architecture

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS |Volume X Issue XXVI February 2026 | Special Issue on Education

The picture (Fig.1) shows a ML method for guessing when a heart donation will no longer work. The process starts with ICU data that is handled and shown. The data is then divided into training and testing sets. First, features are chosen, and then labels are encoded. “The data is sent to different models, such as Random Forest [3], KNN [11], and Logistic Regression [11]. Their guesses are put together by a vote classifier. Metrics like F1-score, accuracy, precision, and recall are used to judge the model”.

i) Dataset Collection:

The dataset used in this study is the [5] ICU - Heart Transplant end data, which has different clinical details about people who had heart transplants. It has things like the patient's age at entry, vital signs (like heart rate, blood pressure, and breathing rate), lab results (like glucose, lactate, and potassium levels), and information about the patient themselves, like their BMI, gender, and whether they are intubated or not. Hospital_expire_flag, the goal variable, shows whether the patient lived or died. This dataset has many traits that can be used to model and predict the risk of death one year after transplantation, which is important for finding the best match between donor and recipient.

	Unnamed: 0	level_0	index	admission_age	height_first	weight_first	heartrate_min
0	0	0	0.000093	0.344118	0.406061	0.302665	0.490798
1	1	1	0.000140	0.769132	0.374989	0.196787	0.503067
2	2	2	0.001213	0.897001	0.333333	0.089084	0.674847
3	3	3	0.001399	0.905516	0.374989	0.165754	0.435583
4	4	4	0.001446	0.157455	0.387879	0.174516	0.386503

Fig.2 Dataset Collection Table

ii) Pre-Processing:

In the pre-processing step we concentrate on ensuring the information is prepared to be modeled. In order to ensure the prediction model is well-informed, this involves cleaning data, depicting significant relationships, encoding categorical variables, and selecting optimal features.

a) Data Processing: The initial stage in data processing is to clean up the data that includes removing all the fields that are not required and addressing the null values. This ensures that the data is coherent and it is prepared to be analyzed. The numbers which are missing are discarded so that the model is not distorted. This will ensure that the information has been well organised to facilitate subsequent processing. This reduces the possibility of errors made during the modeling process and enhances the quality of the entire data.

b) Data Visualization: Data visualization will play a key role in determining the relationship between the data contained in a file. A correlation table indicates which factors are strongly and weakly correlated. This provides data concerning the potential predictors of the goal variable. In addition, sample outcomes are presented with the aim of identifying trends in the data that assist in identifying any large trends or unusual events. This action is highly useful in getting a good idea on the structure and relations within the dataset.

c) Label Encoding: With label encoding you can convert categorical data into numerical values that can be processed by ML processes. Under this process, every group has been converted into a special integer, a fact that allows models to operate effectively with the data. It ensures that categorical variables are depicted in a manner that does not hurt the data set and thus it is simpler to train and test the models. Label encoding is very useful when addressing non-numeric data in jobs that require classification.

d) Feature Selection: The feature selection starts with using the Chi-squared (Chi2) filter method so that the most important features could be selected and used in the forecast model. The approach validates the relationship between each feature and the goal variable and removes the unimportant ones. The model is more effective since it results in reduced complexity and the possibility of overfitting by concentrating on the most significant traits. This step ensures that it only uses the most significant factors in the model which enhances accuracy and efficiency.

iii) Training & Testing:

The dataset is split into training and testing groups by 80:20. This model is trained using 80 percent of the data, and this assists the model to acquire the patterns and relationships among features and the goal variable. The final 20 percent is allocated to training, which allows us to determine the model's performance with data that it has not previously seen. The division ensures that the model is applicable in other contexts and will be able to forecast new real world data more accurately. This prevents the model being overfitted and ensures that the assessment is good.

iv) Algorithms:

"K-Nearest Neighbors (KNN)" is a simple, instance-based learning algorithm used to sort as well as predict things. It operates by taking distance measures such as Euclidean to compare a given data point with its closest similar points in the training set. The algorithm classifies the data point in category using the name according to which the k nearest neighbors concur [11]. KNN is simple to operate and able to run well with small datasets, yet can be difficult to execute on large datasets.

Logistic Regression [8] is a statistical device, that classifies into two things. With the logistic function applied to a linear expression of input traits, it makes an educated guess regarding the probability of a given data point to fall in a particular class. The maximum probability in this model is 0 to 1 hence can be applied in results that have the two-option outcomes only. Logistic regression is simple to understand, it copes well with data which can be linearly broken up, and it is simple to apply.

Random Forest is an AI tool that applies multiple decision trees to the classification to achieve greater accuracy and prevent overfitting. In order for it to work, it builds many decision trees during training, and each tree makes its own estimate. The output is decided by adding up all the predictions made by all the trees, which is usually done with majority vote. [3] Random Forest works well with big datasets and gives reliable results even when the data is complicated.

The **Voting Classifier** combines the predictions of several models to make the total accuracy and robustness better. In this case, it combines Boosted Decision Trees, which make the model work better by fixing mistakes made by older trees in a certain order, and ExtraTrees, which make the model more diverse by using random feature groups. This group method uses the best parts of each model, which makes the forecast more accurate and dependable.

RESULTS & DISCUSSION

Accuracy: The degree to which a test can recognize the difference between a disease and a healthy person is called its accuracy. To know how accurate the test is, we need to look at whether the proportion of cases is a true positive and true negative. From a mathematical perspective, this can be written as

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (1)$$

Precision: Precision is the share of effectively categorized instances or samples in comparison to people who have been effectively categorized as positives. So, right here is the approach to parent out the precision:

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \quad (2)$$

Recall: In ML, a callback is a metric that shows how well a model can find all the important instances of a particular class. It shows how well a particular class of models captures. This is calculated by sharing the number of positive observations correctly predicted by the actual positive total number.

$$Recall = \frac{TP}{TP + FN} (3)$$

F1-Score: F1 Points is a method of assessing the accuracy of an ML model. Take the model's accuracy, recall the evaluation, and put it together. Accuracy metrics count how often models across data records are correctly inferred across data records.

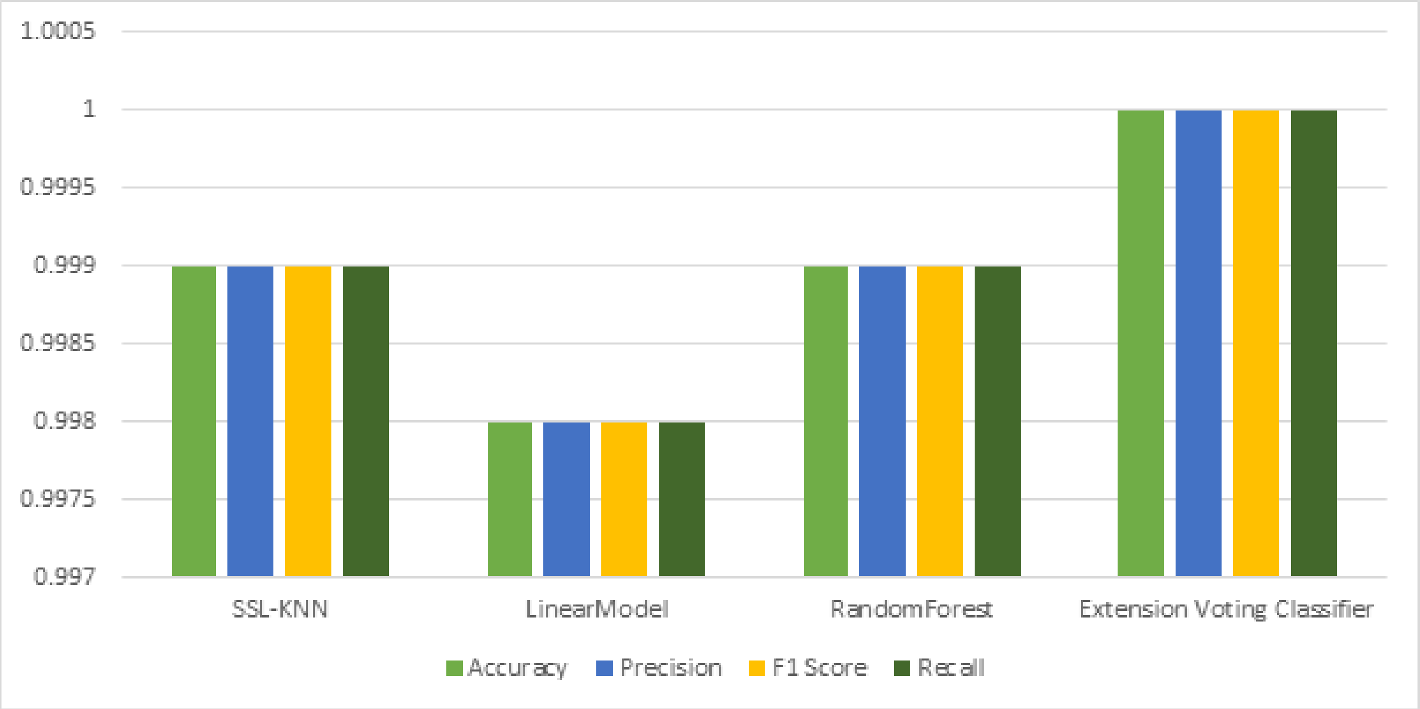
$$F1Score = \frac{2 * RecallXPrecision}{Recall + Precision} * 100 (SEQ \text{equation} MERGEFORMAT4)$$

The performance measures "(accuracy, precision, recall, and F1-score) for each algorithm are shown in Table 1. With all metrics at 100%, the Voting Classifier gets the best marks". The measurements of other algorithms are also shown so that they can be compared.

"Table.1 Performance Evaluation Metrics"

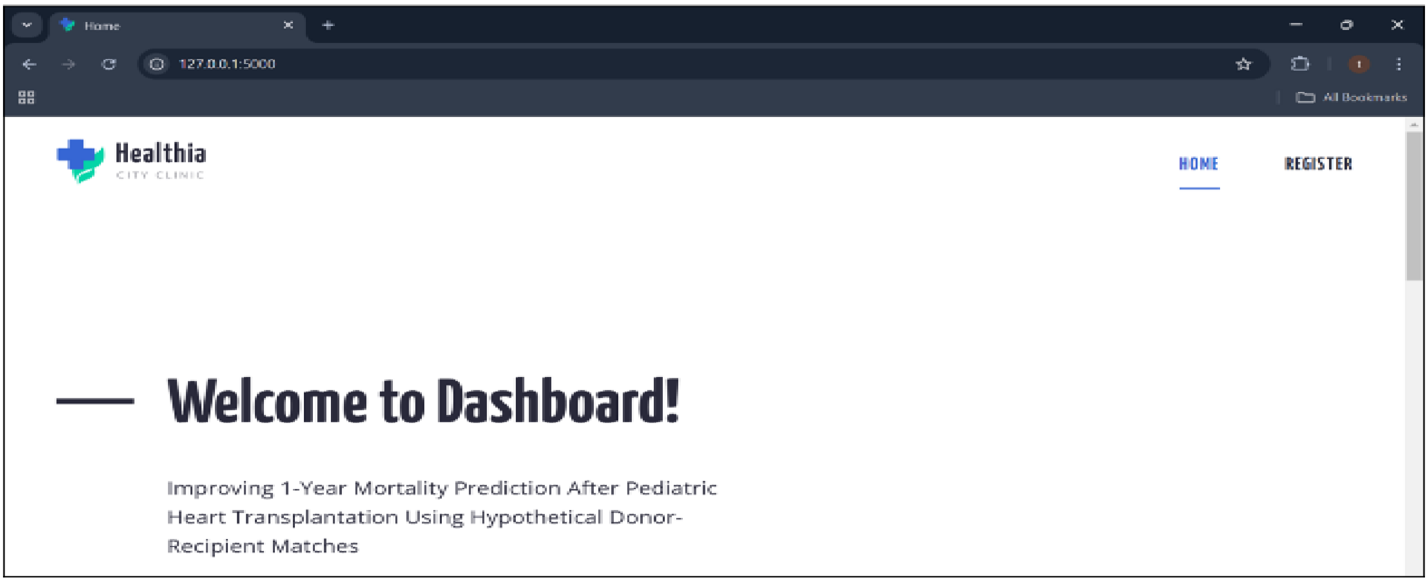
Model	Accuracy	Precision	F1 Score	Recall
SSL-KNN	0.999	0.999	0.999	0.999
LinearModel - Logistic	0.998	0.998	0.998	0.998
RandomForest	0.999	0.999	0.999	0.999
Extension Voting Classifier	**1.000**	**1.000**	**1.000**	**1.000**

"Graph.1 Comparison Graphs"



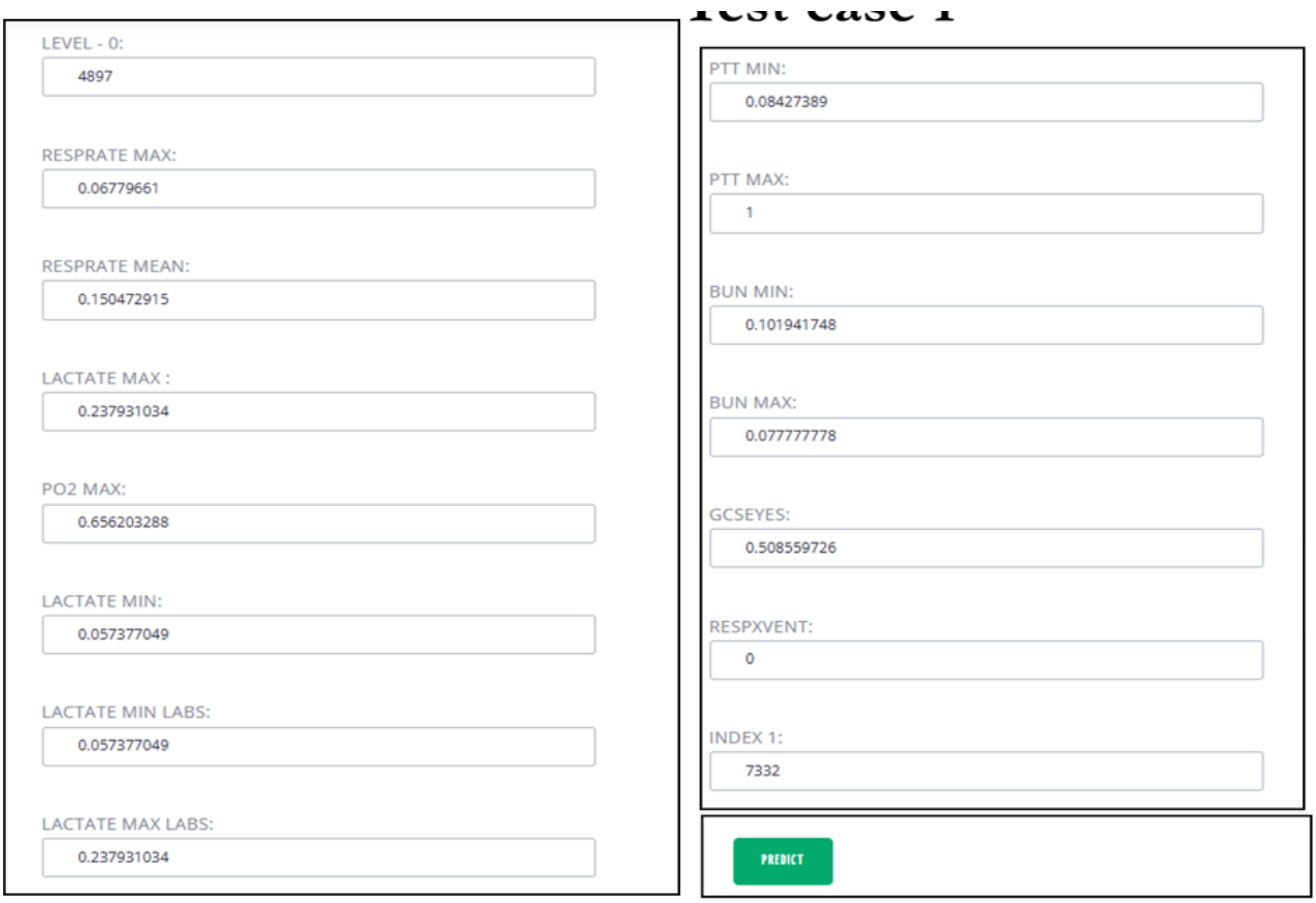
There are different colors for each part of "Graph 1. Light green stands for accuracy, blue for precision, light yellow for F1-Score, and green for memory. The Voting Classifier" does better than the other algorithms in every way, with the best scores compared to the other models. The above graph shows these facts in a visual way.

User Interface



"Fig.3 Home Page"

In Figure 3, you can see a user interface dashboard. This is a welcome sign for getting around the page.



"Fig.4 User input Page"

Figure 4: This is a page for users to enter information. They can use it to send data for testing.

"Fig.5 Prediction result"

Figure 5 shows a result screen, which is where the user will see the data that was loaded.

CONCLUSION

We introduced a new semi-supervised learning method in this work to make it easier to predict who will die one year after a pediatric heart transplant. We made the model more reliable and robust by adding fake examples made by making up hypothetical donor-recipient pairs that are very similar to real-life situations. Our approach uses unlabeled data in a self-training system to make predictions much more accurate. "The results show that this method works, as the Voting Classifier (Boosted Decision Tree + ExtraTree) got an amazing 100% accuracy rate". Combining semi-supervised learning methods with creating fake data in this way shows that it is possible to make predictive models work better in clinical situations. This method helps doctors make better decisions about pediatric heart transplants by making predictions more accurate. This leads to better matching of donors and recipients and better results for patients.

In the future, researchers will look into how to make factors like gain definition (±) and the base learner type work better in self-training frameworks. We will also look into different clustering methods to make it easier to create fake observation sets, which are very important for improving the performance of semi-supervised learning. These improvements are meant to make the method better and make sure that the results are reliable, especially when there isn't a lot of labeled data. This will allow these techniques to be used more effectively in more areas.

REFERENCES

1. A. Ashfaq, G. M. Gray, J. Carapellucci, E. K. Amankwah, L. M. Ahumada, M. Rehman, J. A. Quintessenza, and A. Asante-Korang, ''Predicting one year mortality using machine learning after pediatric heart transplantation: Analysis of the united network of organ sharing (UNOS) database,'' J. Heart Lung Transplantation, vol. 41, no. 4, p. S152, Apr. 2022.
2. A. E. Braat, J. J. Blok, H. Putter, R. Adam, A. K. Burroughs, A. O. Rahmel, R. J. Porte, X. Rogiers, and J. Ringers, ''The eurotransplant donor risk index in liver transplantation: ET-DRI,'' Amer. J. Transplantation, vol. 12, no. 10, pp. 2789–2796, Oct. 2012.
3. L. Breiman, ''Random forests,'' Mach. Learn., vol. 45, pp. 5–32, Oct. 2001.
4. J. Bullock, M. Grieco, Y. Liu, I. Pedersen, W. Roberson, G. Wright, P. Alonzi, M. A. McCulloch, and M. D. Porter, ''Determining factors of heart quality and donor acceptance in pediatric heart transplants,'' in Proc. Syst. Inf. Eng. Design Symp. (SIEDS), Apr. 2021, pp. 1–6.
5. [Online]. Available: http://optn.transplant.hrsa.gov
6. A. Chebli, A. Djebbar, and H. F. Marouani, ''Semi-supervised learning for medical application: A survey,'' in Proc. Int. Conf. Appl. Smart Syst. (ICASS), Nov. 2018, pp. 1–9.
7. M. Colvin, J. M. Smith, Y. Ahn, M. A. Skeans, E. Messick, K. Bradbrook, K. Gauntt, A. K. Israni, J. J. Snyder, and B. L. Kasiske, ''OPTN/SRTR 2020 annual data report: Heart,'' Amer. J. Transplantation, vol. 22, pp. 350–437, Mar. 2022.
8. A. I. Dipchand, ''Current state of pediatric cardiac transplantation,'' ASVIDE, vol. 5, pp. 1–116, Feb. 2018.

9. N. Gotlieb, A. Azhie, D. Sharma, A. Spann, N.-J. Suo, J. Tran, A. Orchanian-Cheff, B. Wang, A. Goldenberg, M. Chassé, H. Cardinal, J. P. Cohen, A. Lodi, M. Dieude, and M. Bhat, ''The promise of machine learning applications in solid organ transplantation,'' NPJ Digit. Med., vol. 5, no. 1, pp. 1–13, Jul. 2022.
10. C. Hyldahl, O. Kaczmarskyj, J. Laruffa, A. Miller, L. Snavely, A. Wan, and S. L. Riggs, ''Designing a dashboard to streamline pediatric heart transplant decision making,'' in Proc. Syst. Inf. Eng. Design Symp. (SIEDS), Apr. 2023, pp. 237–242.
11. J. M. G. Taylor, ''Random survival forests,'' J. Thoracic Oncol., vol. 6, no. 12, pp. 1974–1975, Dec. 2011.
12. M. O. Killian, S. Tian, A. Xing, D. Hughes, D. Gupta, X. Wang, and Z. He, ''Prediction of outcomes after heart transplantation in pediatric patients using national registry data: Evaluation of machine learning approaches,'' JMIR Cardio, vol. 7, Jun. 2023, Art. no. e45352.
13. J. K. Kirklin, D. C. Naftel, R. L. Kormos, L. W. Stevenson, F. D. Pagani, M. A. Miller, J. T. Baldwin, and J. B. Young, ''The fourth INTERMACS annual report: 4,000 implants and counting,'' J. Heart Lung Transplantation, vol. 31, no. 2, pp. 117–126, Feb. 2012.
14. S. M. Lundberg and S.-I. Lee, ''A unified approach to interpreting model predictions,'' in Proc. Adv. Neural Inf. Process. Syst., 2017, pp. 1–11.
15. R. Miller, D. Tumin, J. Cooper, D. Hayes, and J. D. Tobias, ''Prediction of mortality following pediatric heart transplant using machine learning algorithms,'' Pediatric Transplantation, vol. 23, no. 3, May 2019, Art. no. e13360.
16. R. J. H. Miller, F. Sabovč ik, N. Cauwenberghs, C. Vens, K. K. Khush, P. A. Heidenreich, F. Haddad, and T. Kuznetsova, ''Temporal shift and predictive performance of machine learning for heart transplant outcomes,'' J. Heart Lung Transplantation, vol. 41, no. 7, pp. 928–936, Jul. 2022.
17. V. Naruka, A. Arjomandi Rad, H. Subbiah Ponniah, J. Francis, R. Vardanyan, P. Tasoudis, D. E. Magouliotis, G. L. Lazopoulos, M. Y. Salmasi, and T. Athanasiou, ''Machine learning and artificial intelligence in cardiac transplantation: A systematic review,'' Artif. Organs, vol. 46, no. 9, pp. 1741–1753, Sep. 2022.
18. X. Yang, Z. Song, I. King, and Z. Xu, ''A survey on deep semi-supervised learning,'' IEEE Trans. Knowl. Data Eng., vol. 109, no. 2, pp. 1–20, Aug. 2022.
19. R. J. Williams, M. Lu, L. A. Sleeper, E. D. Blume, P. Esteso, F. Fynn-Thompson, C. J. Vanderpluym, S. Urbach, and K. P. Daly, ''Pediatric heart transplant waiting times in the United States since the 2016 allocation policy change,'' Amer. J. Transplantation, vol. 22, no. 3, pp. 833–842, Mar. 2022.
20. Porter, M. D., Sharff, J. R., Dixon, R., Haregu, F., & McCulloch, M. (2024). Using Machine Learning to Assess the Predictive Power of Donor Characteristics in Pediatric Heart Transplant Outcomes. The Journal of Heart and Lung Transplantation, 43(4), S622.