Improving 1-Year Mortality Prediction after Pediatric Heart Transplantation Using Hypothetical Donor-Recipient Matches

Nandgowli Akhil, Palle Avidesh, Sundale Jhansi, Mussam Rajesh Reddy, Nalla Chaitanya

Department of Computer Science Joginpally B.R Engineering College
Department of Computer Science Joginpally B.R Engineering College
Department of Computer Science Joginpally B.R Engineering College
Department of Computer Science Joginpally B.R Engineering College
Department of Computer science Joginpally B.R Engineering College

---***---

Abstract - *The most significant thing that can be done with kids who have end-stage heart failure is heart surgery, though there remains a huge issue of death one year after the transplant. This is a highly significant calculation of this mortality risk to be precise so that the donors and recipients can be better matched and patient outcomes enhanced. In this piece of work, we apply the ICU heart transplant expiration dataset to the determination of the risk of death of pediatric heart transplant patients after one year. We propose an innovative approach that involves methods of advanced feature selection and group to generate more precise predictions. The method includes using Chi-squared tests to select the most important traits and use more than one classifier to make the correct predictions. 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 is a fast and precise technique of estimating the probability of mortality within one year. It provides physicians with valuable data to enhance patient treatment and the most appropriate fit between the donor and recipient during pediatric heart transplants.*

Key Words - Machine learning algorithms, deep learning, classification, sleep disorder, Voting algorithm".

I. INTRODUCTION

Heart transplantation (HTx) has become a procedure that can help to save the lives of children with serious heart failure. Although they constitute approximately 10% of the total number of heart transplants performed annually, there has been a gradual increase over recent decades in the number of cases of pediatric HTx. Over 450 of these surgeries will be performed in United States alone in the year 2020. This has increased with the advancement in medical technology and surgical procedures. However, there are still problems, especially when it comes to lowering the death rate one year after transplantation, which is still very high [7]. Even more difficult, there are not many good organs that can be donated to support pediatric HTx. This adds to the serious problem of people dying while they are on the waiting list. Many pediatric heart donors get discarded as it is difficult to determine whether the organs are quality and whether the recipients will match [7]. That is an indication of the necessity to improve the strategies of donation utilization.

To achieve improved outcomes in pediatric HTx, individuals have been seeking to understand what makes a transplant successful and develop instruments of data visualization to aid physicians to arrive at a decision. Despite all these efforts, the process of matching donors and recipients remains highly subjective and relies on numerous various factors on both sides including medical, physiological and demographic factors [4, 10]. Therefore, to enhance the systems of organ allocation and aid doctors to make improved decisions, it is needed to create reliable prediction models to look at what happens after a transplant [19].

The models of prediction have been of great assistance when making a decision regarding heart transplants. A step that is commonly applied in the allocation process is the HTSS, created by the UNOS in the US. The HTSS considers factors such as the age of the individual, the illness, the degree of functionality and any other health issues that the individual might have such as diabetes or kidney disease. It further examines other aspects of the donor such as their age, cause of death and compatibility of their blood type with the recipient. This score assigns a numerical value to the probability of survival following transplant and assists donation centers to determine the type of patients on waiting lists that it should assist first [13].

Another tool that is established by the Eurotransplant International Foundation is the Eurotransplant Donor Risk Index (ET-DRI), which examines both the factors of the donor and the person receiving the transplant to help them make a choice on whether they should have one. Predictive analytics have been demonstrated to be significant in heart surgery, based on models such as HTSS and ET-DRI. These models aren't perfect, though, because they might not take into account all the factors

that affect how well a transplant works. To enhance the survival rate of recipients, better utilization of organs, and addressing the issues that continue to emerge in pediatric heart transplantation [2], estimates using current data-driven devices must be more precise.
Machine learning can help us make the right judgement about the merits of the donors in kidney heart transplants. This simplifies a lot the prediction of the outcome of the transplant. This approach is a solution to the issue that arises in organ donation as it makes the process more effective and increases the number of patients who survive [20].

II. RELATED WORK

Heart transplantation has emerged as a life-saving procedure to children with end-stage heart failure, yet there are still issues such as allocation of organs and post-transplant death. These issues have attracted the attention of ML methods since they can assist a doctor to make decisions and achieve better outcomes by offering predictive models.

Ashfaq et al. [1] used ML to look at the UNOS database and predict the death rate one year after a pediatric heart donation. They found that ML models could potentially be more effective at outcome prediction compared to conventional statistical approaches because of the fact that they consider a large number of donor and recipient variables. Miller et al. [15] also employed ML techniques including the use of RFs and Neural Networks to enhance prediction of death in pediatric heart transplantation. They demonstrated that ML systems were capable of making more accurate predictions compared to conventional scoring systems. This implies that we ought to abandon the traditional models to use data-driven models to assist doctors make decisions.

Killian et al. [12] examined the national registry data to speculate the outcome of kids who received heart transplants. Their analysis demonstrated the significance of preprocessing data, selecting the appropriate features, and tuning the model to achieve the correct predictions by comparing the methods of ML. The paper also indicated that ML models are capable of adjusting to new data at a fast rate thus suitable to alter clinical circumstances. Gotlieb et al. [9] expanded on this concept and discussed the potential of having ML in solid organ transplants, including heart transplantation. They examined the ways machine learning models can be used to assist in patient selection, organ matching, and postoperative care. They identified certain mechanisms through which ML can reduce the variability in clinical decision-making and enhance outcomes.
Chebli et al. [6] investigated the application of semi-supervised learning to medicine, and how this can be applied to heart transplants. Their solution addressed the issue of unlabeled data scarcity in healthcare through semi-supervised learning methods that provide a good teacher to prediction models. It is particularly effective when doing kid heart transplants as access to the data is difficult. The more the models are able to discover meaningful patterns in both labeled and unlabeled data, the higher the chances that they will be consistent and applicable in the real world.

Miller et al. [16] looked into how the accuracy of ML models in predicting heart transplant results changes over time. They claimed that overtime, ML models become incapable of forecasting the future due to the changing clinical practices, patient populations and organ supply. Their research implied that models have to be continuously re-trained with new information to ensure that the performance remains high. This finding is highly significant to the heart transplants that are performed in the pediatrics, as the rapid advancements in healthcare systems and alteration in the mode of giving out the organs would render the older ones useless.

A single method to interpret statements made by ML models is SHAP (SHapley Additive Explanations), developed by Lundberg and Lee [14]. This approach is particularly effective in medical contexts, such as pediatric heart transplants, where it is useful to doctors to understand how various characteristics can influence the outcomes so that they can make good decisions. SHAP assists physicians in comprehending the manner in which ML predictions were produced by providing them with easy to interpret and consistent feature importance values. This develops trust and openness. The method is widely used in areas where interpretability is important. It gets rid of one of the main problems that makes it hard to use complex ML models successfully in clinical practice.

Naruka et al. [17] did a systematic study that showed how ML and AI can be used in heart transplants. They conducted a study examining the various ML techniques, such as supervised learning and unsupervised learning, and their usefulness in predicting the outcome of transplants, finding the optimal match between the donor and recipient, and determining the quality of organs. The authors emphasized that ML can correct the issues of existing allocation techniques and improve in the long-run. Nevertheless, they also identified such issues as the standardization of data and ethical issues. In order to reap the full benefits of ML in the area of cardiac transplantation, they urged physicians and data scientists to collaborate.
Yang et al. [18] conducted an extensive review of deep semi-supervised learning, and the authors are interested in comprehending how it could be applied in the case of

limited amounts of labeled data, such as infant heart transplantation. They discussed new techniques, including generative models and consistency regularization, which can improve the training on both labeled and unlabeled data. This is particularly effective in the medical environment where tagged data is difficult to locate and expensive to access. Deep semi-supervised learning has the potential to enhance performance and generalization of a model by using unlabeled data well. This renders it a promising approach towards enhancing ML applications in healthcare.

III. MATERIALS AND METHODS

Our method is more sophisticated and consists of a new combination of ML and feature selection to predict whether a child will die one year after heart transplantation or not. The ICU heart transplant expiration dataset will be used by the system to find important factors that raise the risk of death after donation. Chi-squared (Chi2) method will be used to weed out irrelevant traits and ensure the model concentrates on the most significant predictors. This will enhance our predictions. To make predictions, we will then apply some ML techniques, including KNN [11], LR [8], and RF [3]. We shall also consider ensemble techniques particularly the Voting Classifier, which combines the strengths of the Boosted Decision Trees and ExtraTree in order to make the system perform better through the combination of their strengths. A big part of the suggested system will be using semi-supervised learning methods. The falsification of cases in these techniques consists of pairing donors and recipients in a manner that resembles real cases greatly. This will allow the system to utilize the data that has not been labeled and thus will be more accurate in making correct guesses. The proposed approach will enhance clinical outcomes and ensure easier finding of each other by the donors and recipients.

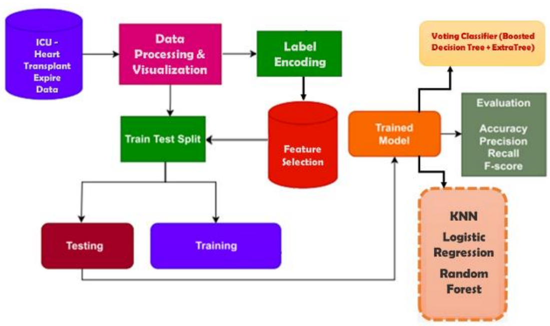

Fig.-1: Proposed Architecture

In the image (Fig.1), there is a ML approach to predicting that a heart donation will not be effective any more. This begins with ICU data, which is processed and presented. The data is subsequently split into training and testing set. Features are first selected, then encoded. The data is sent to different models, such as RF [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 data used in this research is the [5] ICU - Heart Transplant end data, which contains varying clinical information of individuals that underwent heart transplants. It includes such information as the age of the patient on admission, vital signs (heart rate, blood pressure, breathing rate), laboratory findings (glucose, lactate and potassium levels), and data about the patient him/herself, such as BMI, gender, and intubation or non-intubation. The goal variable is hospital_expire_flag, which displays the survivability/death of the patient. 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 stage, we focus on preparing the dataset to model. This includes data cleaning, visualization of meaningful relationships, coding nominal values, and use of feature selection to ensure the model has the best input.

a) Data Processing: The purification of the dataset, which includes the removal of extraneous columns and the null values, is the first part of the data processing stage, which guarantees the data consistency and readiness to be analyzed. Any missing values are removed to prevent distortions in the model. This step ensures that the dataset is structured in such a way that it is further manageable with minimal likelihood of errors in the modeling process and a higher quality of data in general.

b) Data Visualization: To understand how the data is related to each other, you need to be able to see it. A correlation table has strong and weak links between features and this provides you with an impression of what variables can predict the goal variable. Results of a sample are also plotted in a manner that allows comparison of data trends. This makes it easier to see important patterns or outliers. This is a very important step to be able to tell how the dataset is structured and how it is interrelated.

c) Label Encoding: It is called label encoding and is a method of storing categorical data using numbers, and this makes it ideal in machine learning techniques. This step assigns unique numbers to each group, enabling the data to be viewed using models. It makes sure that categorical factors are shown in a way that doesn't damage the dataset, which makes training and testing models better. The label encoding is particularly applicable in non-number data in classification tasks.

d) Feature Selection: This method uses the Chi2 filter to pick out the most important traits that will be used in the predictive model. This method figures out how each attribute is related to the goal variable and then gets rid of the ones that aren't important. By concentrating on the most significant features, the model will perform better, become easier, and reduce the chances of overfitting. This measure will ensure that the model is influenced by the most significant factors only which will make it more accurate and useful.

iii) Training & Testing:

Two-fifths of the data is training and the other two-fifths are testing. Training the model consumes eighty percent of the data and assists it when discovering the fundamental patterns and correlation of the characteristics with the objective variable. One can observe the effectiveness of the model with respect to the bizarre data since the remaining 20% are the case tests. This division ensures that the model is applicable to other scenarios and capable of estimating new real world of data. This prevents the model being over fitted and makes it open to healthy appraisal.

iv) Algorithms:

This is a basic instance based learning algorithm known as K-Nearest Neighbors (KNN) which is employed to address classification and regression problems. It determines what to do based on distance measures such as Euclidean distance to make comparisons of a piece of data with similar pieces in the training set. It classifies the data points into clusters according to the label that is most popular as per the KNN [11]. KNN is also simple to comprehend and can be used on small datasets. Nevertheless, it may be difficult to work with large datasets since it consumes a lot of computing power.

Logistic Regression is a statistical modeling to address two-choice problems. It approximates the probability of a piece of data in a given group using the logistic function on a linear combination of factors of input. The model assigns probabilities in between 0 to 1 and thus can be applied to both true and false results. Logistic regression is straightforward, understandable, and applicable when there is a straight line that the data can be divided into.

The simple linear regression line,

^y=a+bx y ^ = a + b x (1)

can be taken to mean:

Where y is the expected value of y, a is the intercept (where the regression line crossing the y-axis), and b is the change in y per unit change in x.

Random Forest is an ensemble algorithm, which combines many decision trees to increase the accuracy of classification and reduce overfitting. It works on the principle of creating multiple decision trees in the training stage, where each tree is independent in its forecast. The ultimate output is ascertained by consolidating the forecasts of all trees, generally through majority vote. 3 RF is efficient in working with large datasets and providing high performance, particularly in working with complex data.

The **Voting Classifier** is a composite of many models that are used to enhance the overall accuracy and strength. This method uses both Boosted DT that improve the performance of the model by progressively correcting the errors of the previous trees, and ExtraTrees that increase the diversity of the model by using random feature subsets. Such an ensemble approach enjoys the benefits of various models, providing increased accuracy and reliability in predictions.

IV. RESULTS AND DISCUSSION

Accuracy: Accuracy of a test is the ability of a test to distinguish between patients and healthy individuals. In order to measure the accuracy of tests, calculate the proportions of true positive and true negative results of all the instances that are tested. This would be mathematically expressed as:

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \quad (2)$$

Precision: Precision is a measure of the percentage of recognised positive cases or samples. The formula used to determine precision is:

$$Precision = \frac{\text{True Positive}}{\text{True Positive } + \text{False Positive}} \quad (3)$$

Recall:ML recall is the measure of how well a model determines all relevant examples of a category. It illustrates the effectiveness of a model in summarizing the cases of a class by accurately relating the expected positive cases with the number of positives in total.

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \quad (4)$$

F1-Score: The F1 score is used to determine the accuracy of a ML model. Combining model accuracy and recall. The accuracy measure determines the number of accurate predictions that a model makes over a dataset.

$$F1\ Score = 2 * \frac{Recall\ \mathrm{X}\ Precision}{Recall + Precision} * 100 \quad (5)$$

Table 1 shows the performance measures of accuracy, precision, recall and F1-score evaluated against each algorithm. The Voting Classifier scores best, with all at 100%. Alternative method metrics are also given to compare.

Table.1: Performance Evaluation Metrics of classification

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**

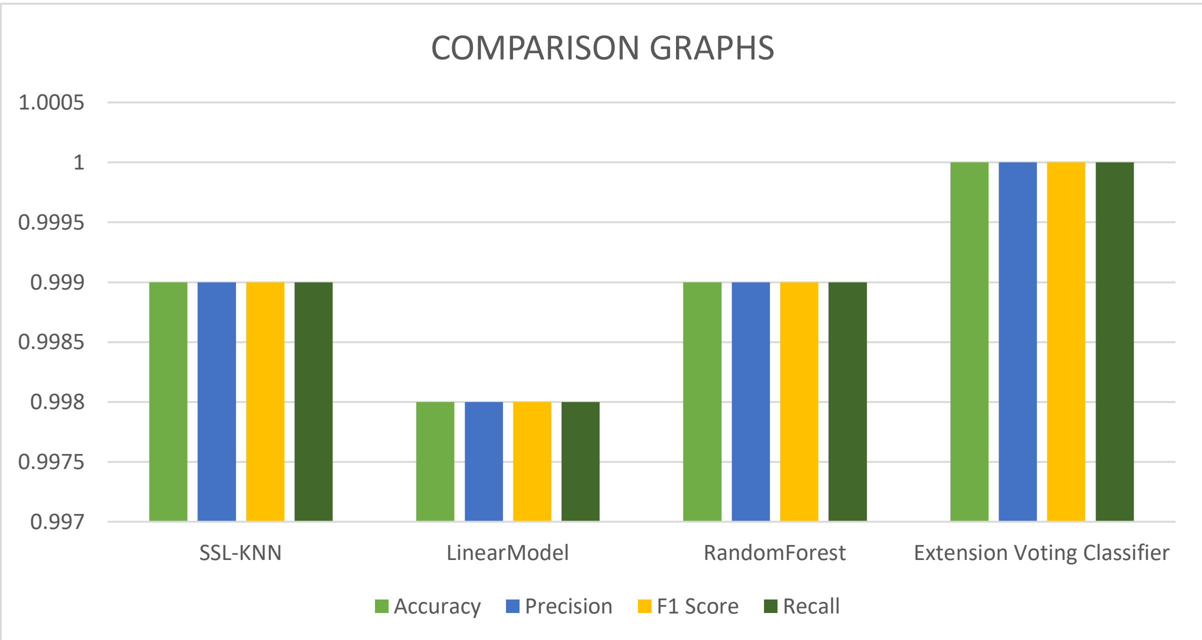

Fig-3: Comparison Graphs of Classification

Accuracy, precision, F1-Score, and recall are represented in light green, blue, light yellow, and green, respectively in Graph 1. The Voting Classifier outperforms all the other algorithms in all settings, and has the best values compared to the other models. These features can be graphically displayed in the graph above.

V. CONCLUSION

The research is a distinctive semi-supervised type of learning that focuses on improving the accuracy of the one-year death predictions after the heart transplantation in children. To enhance the reliability and stability of the model, we incorporated synthetic examples that are designed based on hypothetic donor-recipient pairings that are as close to real-world scenarios as possible. Our algorithm involves the use of unlabeled data within a self-training system, which significantly boosts prediction accuracy. The results confirm the effectiveness of this approach, as the Voting Classifier (Boosted Decision Tree + ExtraTree) has achieved an impressive accuracy of 100%. This algorithm shows the effectiveness of semi-supervised learning coupled with synthetic data to improve predictive models in clinical settings. This plan will improve the prediction accuracy, which is of great help in decision-making in the field of pediatric heart transplantation, therefore, the optimal matching of the donor and recipient and the improved outcomes of patients.

Future studies will explore how to optimize parameters such as the definition of gain ( ) and the kind of base learner in self-training models. Additionally, novel clustering algorithms will be investigated to enhance the production of synthetic observation sets, which are crucial in enhancing semi-supervised learning effectiveness. These advances aim at improving the methodology and ensuring good results particularly in cases where there is limited labelled data so that the techniques can be more effectively applied in a wide range of disciplines.

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.