Machine Learning Techniques for Donor-Recipient Matching in Liver Transplantation

DePaul University

MS Data Science

Chien Lin Yang

Outline

1. Background & Purpose
2. Data
3. Model
4. Results & Findings
5. Conclusion

Background

- Donor-Recipient (D-R) matching is a challenging topic in Liver Transplantation (LT) due to the increasing number of candidates for LT and the limited number of available donors.
- Some scores have been designed to aid this process, such as the Model for End-Stage Liver Disease (MELD). The primary goal of these scores is to decrease mortality in the waiting list without affecting the result of the transplant.
- However, in some cases, a decrease in waiting list mortality leads to worse post-transplant survival results.

Purpose

- Try to use the machine learning models to predict graft survival at different time points after transplantation.
- Graft survival:
  - ✓ 3 months (3M)
  - ✓ 1 year (1M)
  - ✓ 2 years (2M)
  - ✓ 5 years (5M)

Data

A real-world data

- UNOS: United Network for Organ Sharing (UNOS) is the non-profit serving as the nation’s transplant system under contract with the federal government.
- Liver data
- Samples: 332,787
- Variables: 426

Model

1. Logistic Regression (LR)
2. Decision Tree (DT)
3. Random Forest (RF)
4. Gradient Boosting (GB)
5. K-nearest neighbors (KNN)
6. Support Vector Machines (SVM)

- Evaluation
  1. Confusion Matrix (CM)
  2. Accuracy (Acc)

Data Processing

- Simulate the data processing process of this paper[1] to build different machine learning models for binary classification problems with different end-points.

1. Guijo-Rubio D, Briceño J, Gutiérrez PA, Ayllón MD, Ciria R, Hervás-Martínez C. Statistical methods versus machine learning techniques for donor-recipient matching in liver transplantation. PLoS One. 2021 May 21;16(5):e0252068. doi: 10.1371/journal.pone.0252068. PMID: 34019601; PMCID: PMC8139468.

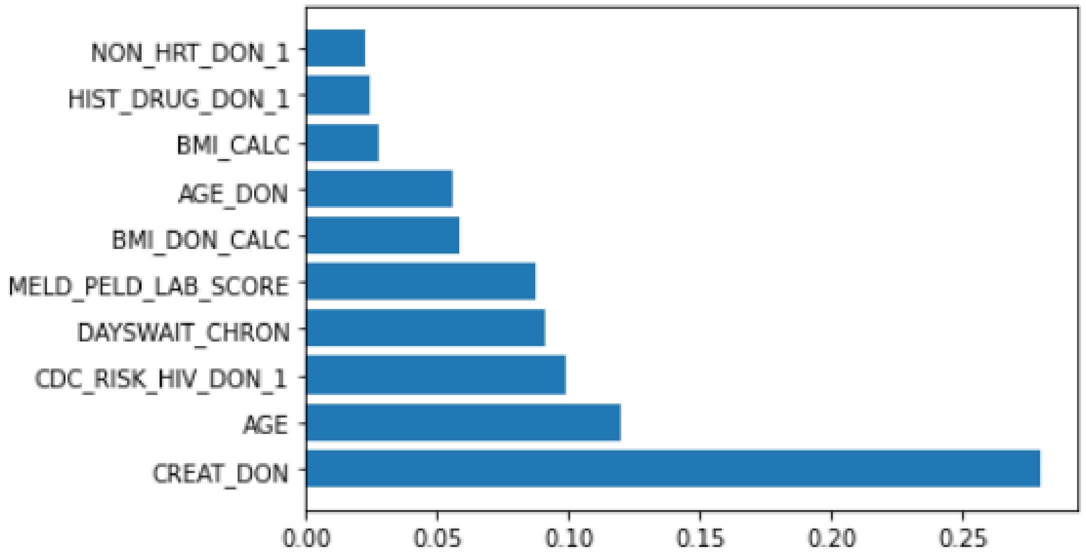
Results

- Among 3M, 1Y, and 2Y end-points, all the models obtain similar results.
- Accuracy: 95% at 3M; 84% at 1Y; 72% at 2Y
- Focusing on the 5Y end-point, the Gradient Boosting (GB) has the best performance in the test dataset.

End-Point	LR	DT	RF	GB	KNN	SVM
3M	0.9510	0.9510	0.9510	0.9510	0.9495	0.9510
1Y	0.8399	0.8399	0.8399	0.8399	0.8188	0.8399
2Y	0.7238	0.7238	0.7240	0.7254	0.6877	0.7239
5Y	0.6225	0.5368	0.6361	$\boxed{0.6780}$	0.5726	0.5525

Findings

- The Gradient Boosting (GB)
- Top 10 Feature Importances
- The important variables are the index of deceased donor terminal lab creatinine **CREAT_DON ,** the age of recipient and donor **AGE** and **AGE_DON**, among others.



Conclusion

- In this project, the best results are obtained by the Gradient Boosting (GB) method, especially in 5 years end-point.
- However, the performance of each model is similar among the other end-points, and one of the reasons behind this is the unbalanced class in each end-point, although I have used Recall in the scoring of the confusion matrix. In the future, I will try to improve the unbalanced data so that I can get better results.

Thank you!