Lung Transplant Outcome Prediction using UNOS Data

Ankit Agrawal*, Reda Al-Bahrani*, Mark J. Russo†, Jaishankar Raman‡ and Alok Choudhary*
*Dept. of Electrical Engineering and Computer Science,
Northwestern University,
Evanston, IL USA
Email: {ankitag,rav650,choudhar}@eecs.northwestern.edu
†Department of Cardiothoracic Surgery,
Barnabas Health Heart Centers
Livingston, NJ USA
Email: lmrusso@barnabashealth.org
‡Department of Cardiothoracic & Vascular Surgery,
Rush University Medical Center
Chicago, IL USA
Email: jai_raman@rush.edu

***Abstract*—We analyze lung transplant data from the United Network for Organ Sharing (UNOS) program with the aim of developing accurate risk prediction models for mortality within 1 year of lung transplant using data mining techniques. The data used in this study is de-identified and consists of 62 predictor attributes, and 1-year posttranplant survial outcome for patients who underwent lung transplant between the years 2005 and 2009. Our dataset had 5,319 such patient instances. Several data mining classification techniques were used on this data along with various data mining optimizations and validations to build predictive models for the abovementioned outcome. Prediction results were evaluated using c-statistic metric, and the highest c-statistic obtained was 0.68. Further, we also applied feature selection techniques to reduce the number of attributes in the model from 50 to 8, without any degradation in c-statistic. The final model was also found to outperform logistic regression, which is the most commonly used technique in predictive healthcare informatics. We believe that the resulting predictive model on the reduced dataset can be quite useful to integrate in a risk calculator to aid both physicians and patients in risk assessment.**

***Keywords*-lung transplant; predictive modeling; data mining;**

I. Introduction

A lung transplant, or pulmonary transplant, is a surgical transplant procedure in which a patient's diseased lungs are partially or totally replaced by lungs which come from a donor. [1]. As of 2008, the survival rate for lung transplant after 1 year was 83.6% [2]. Complications of lung transplantation include rejection of the transplanted lung and infection [3]. Typical expenses range from around $600,000 to $1,100,000 for single lung, double lung, and heart-lung transplant [4], [5].

As organs available for transplant remain critically scarce, achieving maximal benefit from lung transplantation depends upon improved recipient and donor selection [6]. Thus accurate estimation of lung transplant outcomes can improve both informed patient consent by helping patients better understand its risks and benefits, and also aid the physicians in decision making by assessing the true patient-specific risks of the operation, rather than relying on population-wide risk assessments. To this end, accurate outcome prediction of performing transplantation is extremely important.

The United Network for Organ Sharing (UNOS) is a private, non-profit organization that manages the nation's organ transplant system under contract with the federal government [7]. UNOS is involved in many aspects of the organ transplant and donation process, including maintaining the database that contains all organ transplant data for every transplant event that occurs in the US.

Applying data mining techniques to lung transplantation data can be useful to rank and link pretransplantation attributes to the outcome. Here we use data mining techniques on UNOS lung transplantation data to estimate 1-year survival of lung transplant patients, based on pretransplant characteristics. Experiments with nearly 50 modeling techniques were conducted and the results compared to find the best model for the data used in this study. It was found that rotation forest ensembles of alternation decision trees resulted in the best discrimination (c-statistic) between survived and non-survived lung recepients. Further, feature selection was used to find a smaller subset of attributes that can potentially achieve similar prediction performance, but can result in a simpler model.

The rest of the paper is organized as follows: Section 2 describes the data mining techniques used in this study followed by a brief description of the UNOS data used in this study in Section 3. Experiments and results are presented in Section 4, and the conclusion and future work is presented in Section 5.

978-1-4799-1293-3/13/$31.00 ©2013 IEEE

II. Data Mining Techniques

A. Modeling

We used 47 classification schemes in this study, including both direct application of classification techniques and also constructing their ensembles using various ensembling techniques. Due to space limitations, here we briefly describe only those classification/ensembling techniques whose results we present in the next section.

1) **Support vector machines**: SVMs are based on the Structural Risk Minimization(SRM) principle from statistical learning theory. A detailed description of SVMs and SRM is available in [8]. In their basic form, SVMs attempt to perform classification by constructing hyperplanes in a multidimensional space that separates the cases of different class labels. It supports both classification and regression tasks and can handle multiple continuous and nominal variables.
2) **Artificial neural networks**: ANNs are networks of interconnected artificial neurons, and are commonly used for non-linear statistical data modeling to model complex relationships between inputs and outputs. The network includes a hidden layer of multiple artificial neurons connected to the inputs and outputs with different edge weights. The internal edge weights are 'learnt' during the training process using techniques like back propagation. Several good descriptions of neural networks are available [9], [10]. It has been used for accurate estimation in different areas such as mobile health [11], drug abuse [12], civil engineering [13], computer vision [14] and video delivery [15], [16].
3) **Decision Table**: Decision table typically constructs rules involving different combinations of attributes, which are selected using an attribute selection search method. Simple decision table majority classifier [17] has been shown to sometimes outperform state-of-the-art classifiers.
4) **KStar**: KStar [18] is a lazy instance-based classifier, i.e., the class of a test instance is based upon the class of those training instances similar to it, as determined by some similarity function. It differs from other instance-based learners in that it uses an entropy-based distance function.
5) **Reduced error pruning tree**: Commonly known as REPTree [19], it is a implementation of a fast decision tree learner, which builds a decision/regression tree using information gain/variance and prunes it using reduced-error pruning.
6) **Random forest**: The Random Forest [20] classifier consists of multiple decision trees. The final class of an instance in a Random Forest is assigned by outputting the class that is the mode of the outputs of individual trees, which can produce robust and accurate classification, and ability to handle a very large number of input variables. It is relatively robust to overfitting and can handle datasets with highly imbalanced class distributions.
7) **Alternating decision tree**: ADTree [21] is decision tree classifier which supports only binary classification. It consists of two types of nodes: decision nodes (specifying a predicate condition, like 'age' $>$ 45) and prediction nodes (containing a single real-value number). ADTrees always have prediction nodes as both root and leaves. An instance is classified by following all paths for which all decision nodes are true and summing the values of any prediction nodes that are traversed. This is different from the J48 decision tree algorithm in which an instance follows only one path through the tree.
8) **Decision stump**: A decision stump [19] is a weak tree-based machine learning model consisting of a single-level decision tree with a categorical or numeric class label. Decision stumps are usually used in ensemble machine learning techniques.
9) **Naive Bayes**: The naive bayes classifier [22] is a simple probabilistic classifier that is based upon the Bayes theorem. This classifier makes strong assumptions about the independence of the input features, which may not always be true. It makes use of the variables contained in the data sample, by observing and relating them individually to the target class, independent of each other. Despite this assumption, the naive bayes classifier works well in practice for a wide variety of datasets and often outperforms other complex classifiers.
10) **Bayesian Network**: A Bayesian network is a graphical model that encodes probabilistic relationships among a set of variables, representing a set of random variables and their conditional dependencies via a directed acyclic graph (DAG). Bayesian network learning can be used with various search algorithms for searching the network structures, and estimator algorithms for finding the conditional probability tables of the network.
11) **Logistic Regression**: Logistic Regression [23] is used for prediction of the probability of occurrence of an event by fitting data to a sigmoidal S-shaped logistic curve. Logistic regression is often used with ridge estimators [24] to improve the parameter estimates and to reduce the error made by further predictions.
12) **AdaBoost**: AdaBoost [25] is a commonly used ensembling technique for boosting a nominal class classifier. In general, boosting can be used to significantly reduce the error of any weak learning algorithm that consistently generates classifiers which need only be a little bit better than random guessing. It usually dramatically improves performance, but is also prone

to overfitting.

13) **LogitBoost**: The LogitBoost algorithm is an ensembling technique implementation of additive logistic regression which performs classification using a regression scheme as the base learner, and can handle multi-class problems. In [26], the authors explain the theoretical connection between Boosting and additive models.
14) **Bagging**: Bagging [27] is a meta-algorithm to improve the stability of classification and regression algorithms by reducing variance. Bagging is usually applied to decision tree models to boost their performance. It involves generating a number of new training sets (called bootstrap modules) from the original set by sampling uniformly with replacement. The bootstrap modules are then used to generate models whose predictions are averaged to generate the final prediction. Bagging has been shown work better with decision trees than with linear models.
15) **Random subspace**: The Random Subspace classifier [28] constructs a decision tree based classifier consisting of multiple trees, which are constructed systematically by pseudo-randomly selecting subsets of features, trying to achieve a balance between overfitting and achieving maximum accuracy. It maintains highest accuracy on training data and improves on generalization accuracy as it grows in complexity.
16) **Rotation Forest**: Rotation forest [29] is a method for generating classifier ensembles based on feature extraction, which can work both with classification and regression base learners. The training data for a the base classifier is created by applying Principal Component Analysis (PCA) [30] to K subsets of the feature set, followed by K axis rotations to form the new features for the base learner, to encourage simultaneously individual accuracy and diversity within the ensemble.

B. Feature Selection

Feature selection techniques are typically used to select a subset of relevant features for use in a model. It is based on the assumption that data contains many redundant or irrelevant attributes that do not add much to the information provided by other existing attributes. We used 2 feature selection techniques in this study:

1) **Correlation Feature Selection (CFS)**: CFS is used to identify a subset of features highly correlated with the class variable and weakly correlated amongst them [31]. CFS was used in conjunction with a greedy stepwise search to find a subset S with best average merit, which is given by:

$$Merit_S = \frac{n.\overline{r_{fo}}}{\sqrt{n + n(n-1).\overline{r_{ff}}}}$$

where n is the number of features in S, $\overline{r_{fo}}$ is the average value of feature-outcome correlations, and $\overline{r_{ff}}$ is the average value of all feature-feature correlations.

2) **Information Gain**: This is used to assess the relative predictive power of the predictor attributes, which evaluates the worth of an attribute by measuring the information gain with respect to the outcome status:

$$IG(Class, Attrib) = H(Class) - H(Class|Attrib)$$

where $H(.)$ denotes the information entropy.

The CFS technique evaluates subsets rather than individual attributes, so it was first used to find a smaller subset of attributes. Subsequently, information gain was used on the reduced set of attributes to get a ranking of the attributes in the order of their predictive potential, as information gain evaluates each attribute independently.

III. UNOS Data

The UNOS STAR Thoracic Organ Transplant and Waiting List File contains data on all transplant candidates and transplant recipients of heart, lung, and heart-lungs that have been listed or performed in the U.S. and reported to the OPTN since October 1987. Data entry by all US transplant centers is mandated by the 1984 National Transplantation Act. This cohort totals over 37,000 observations. There is one record per waiting list registration/transplant event. Each record includes the most recent follow-up data, including patient and graft survival, waittime, and the patient's list status (e.g., waiting, transplanted, removed prior to transplant, or dead). This dataset contains nearly 500 fields to characterize candidate/recipient and donor information including demographics (eg, age, race, sex), social history, and clinical information (eg, blood type, measures of lung function and hemodynamic measures, past medical and surgical history, serologies, and severity of co-morbid illness).

UNOS provided de-identified patient-level data (data source #01052011-6). These data include all lung transplant recipients and donors in the U.S. and reported to the Organ Procurement and Transplantation Network between January 1, 2005 and December 31, 2009. The use of these data is consistent with the regulations of our Institutional Review Board.

The study population included 5,319 lung transpalnt patients aged 18 years and older between January 1, 2005 and December 31, 2009. Patients were monitored from the date of transplantation to January 3, 2011, which was the last day of follow-up provided by UNOS. 62 predictor attributes were assessed. The primary outcome variable was 1-year post-transplant survival. We omit the details of all the input 62 attributes here due to space constraints. A brief description of the selected subset 12 attributes used in the final model is presented later. Out of 5,319 patients, 1,061 patients (19.95%) did not survive more than 1 year after transplant.

IV. Experiments and Results

We used the WEKA toolkit 3.6.7 for the implementation of data mining techniques described earlier [32]. 3-fold cross-validation was used for evaluation. Cross-validation is routinely used to evaluate the prediction performance of data mining models to eliminate any chances of over-fitting. In k-fold cross-validation, the input data is randomly divided into k segments. $k-1$ segments are used to build the model and the remaining 1 segment unseen by the model is used to test/evaluate it. This is repeated k times with different test segments, and the results are aggregrated. In this way, each instance of the dataset is run through a model that has not seen it during the training phase. Running a test instance through a trained model generates a probability distribution for that instance to belong to different possible class values (here, binary 1-year survival). A probability cutoff is required to actually classify the test instance into one of the classes. For binary classification, 50% cutoff is most commonly used.

A. Evaluation metrics

Binary classification performance can be evaluated using various metrics. We use the following in this work:

1) **c-statistic (AUC)**: The ROC (Receiver operating characteristic) curve is a graphical plot of true positive rate and false positive rate. The area under the ROC curve (AUC or c-statistic) is an effective metric for evaluating binary classification performance, as it is independent of the probability cutoff and measures the discrimination power of the model. This is the primary metric used in this work for inter-model comparison.
2) **Overall accuracy**: It is the percentage of predictions that are correct. For highly unbalanced classes where the minority class is the class of interest, overall accuracy by itself may not be a very useful indicator of classification performance, since even a trivial classifier that simply predicts the majority class would give high values of overall accuracy.

$$Overall\ accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

where TP is the number of true positives (hits), TN is number of true negatives (correct rejections), FP is number of false positives (false alarms), and FN is number of false negatives (misses).
3) **Sensitivity (Recall)**: It is the percentage of positive labeled records that were predicted positive. Recall measures the completeness of the positive predictions.

$$Sensitivity = \frac{TP}{(TP + FN)}$$

4) **Specificity**: It is the percentage of negative labeled records that were predicted negative, thus measuring the completeness of the negative predictions.

$$Specificity = \frac{TN}{(TN + FP)}$$

5) **Positive predictive value (Precision)**: It is the percentage of positive predictions that are correct. Precision measures the correctness of positive predictions.

$$Positive\ predictive\ value = \frac{TP}{(TP + FP)}$$

6) **Negative predictive value**: It is the percentage of negative predictions that are correct, thereby measuring the correctness of negative predictions.

$$Negative\ predictive\ value = \frac{TN}{(TN + FN)}$$

7) **F-measure**: It is in general, possible to have either good precision or good recall, at the cost of the other, and F-measure combines the two measures in a single metric by taking the harmonic mean of precision and recall.

$$F - measure = \frac{2.precision.recall}{(precision + recall)}$$

B. Modeling and Feature Selection

As mentioned before, we used 47 classification schemes on this data. Fig. 1 present the results on 15 classification schemes for 1-year survival, consisting of most of the popular classifiers. For each of the ensembling techniques, many underlying classfiers were tried in the experiments but only the one with the best c-statistic is preented in the figure. Blue bars represent the c-statistic with the entire set of 62 attributes, and red bars represent the results with the reduced set after feature selection. Using correlation based feature selection (CFS) technique yielded a subset of only 12 features for the given outcome of 1-year survival.

The figures clearly show that many of the evaluated classification schemes perform comparably well for 1-year survival. Of all the models used in this study, Rotation Forest with Alternate Decision Trees as the underlying classifier gave the best c-statistic of 0.677 with 62 attributes, and of 0.680 with 12 attributes. Thus, feature selection techniques were able to identify a much smaller subset without a loss in c-statistic.

Figure 2 presents the relative predictive power of the resulting smaller subset of attributes identified by CFS for 1-year survival. Following is a brief description of these 12 attributes:

1) **Ability to perform daily activities**: Functional status is an individual's ability to perform normal daily activities required to meet basic needs, fulfill usual roles, and maintain health and well-being. Decline in functional status is measured by an individual's loss of independence in activities of daily living (ADLs) over a period of time [33].

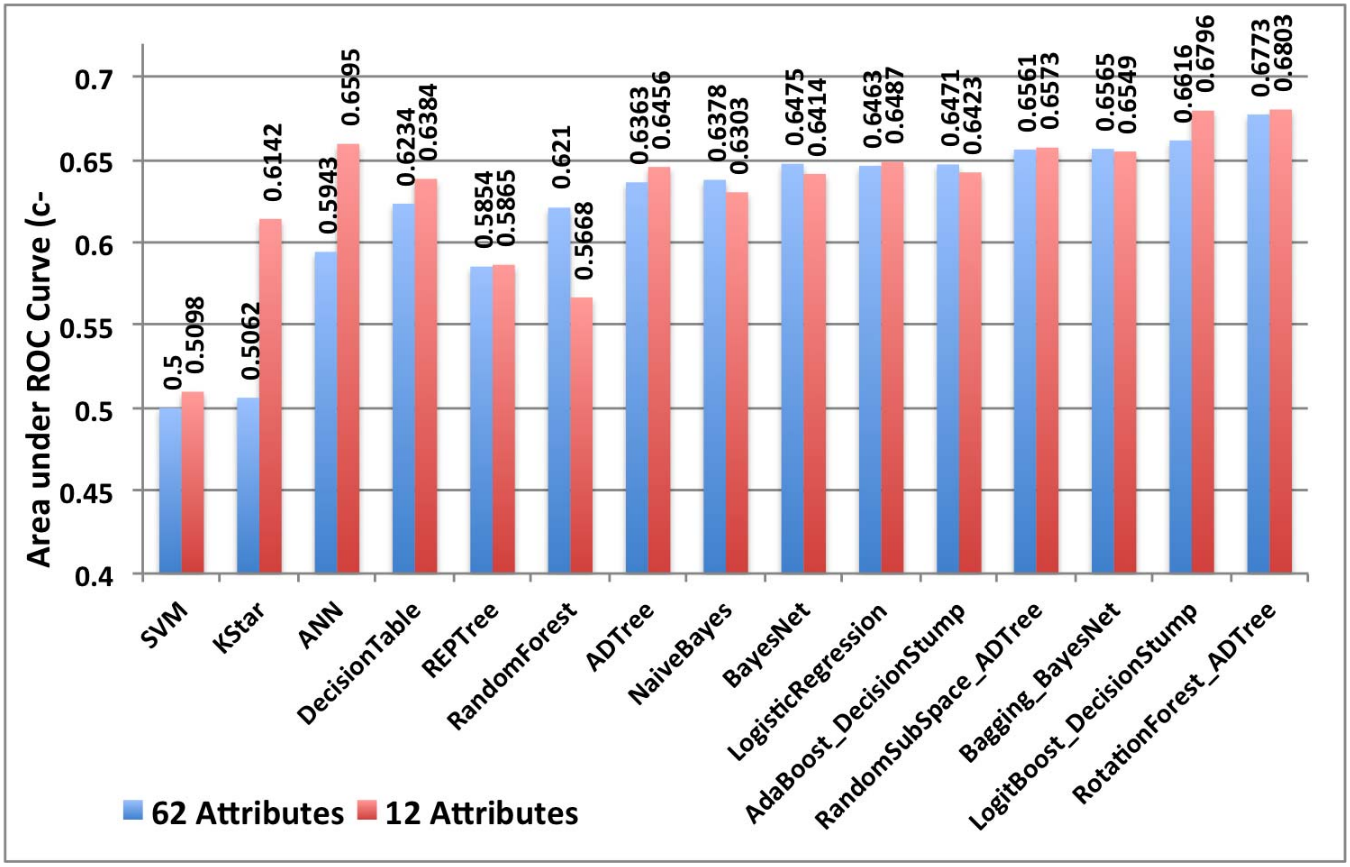

Figure 1. Prediction performance comparison for 1-year survival in terms of area under the ROC curve (c-statistic).

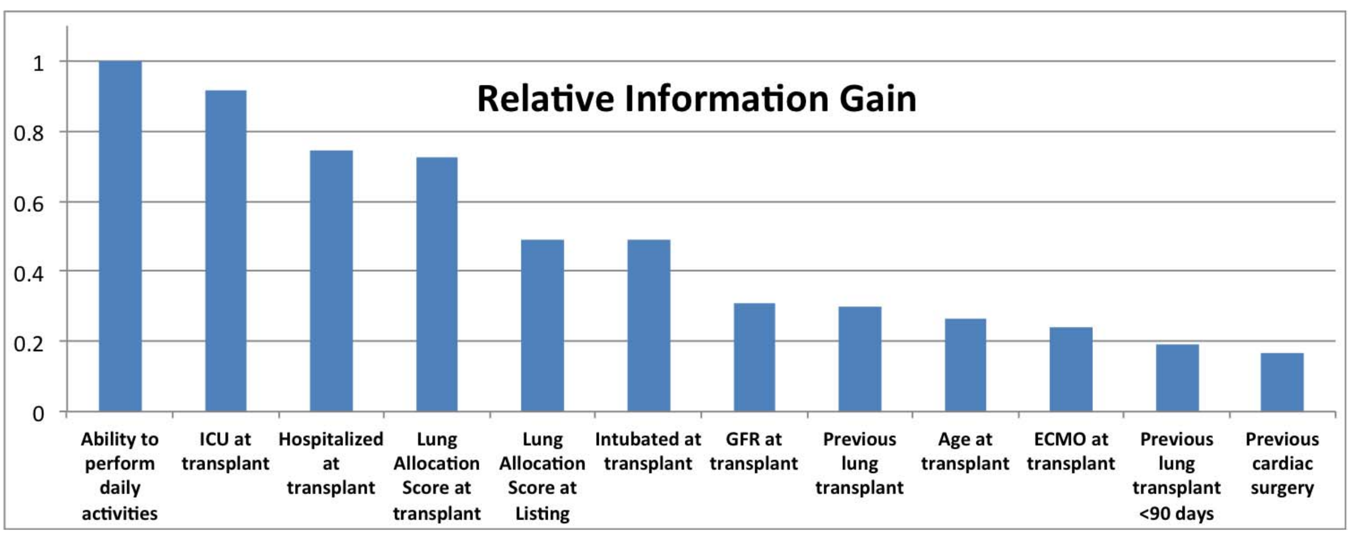

Figure 2. Relative information gain of features resulting from the CFS technique for 1-year survival.

2) **ICU at transplant**: It represents whether the patient was in Intensive Care Unit at the time of transplant. Intensive Care Units cater to patients with the life-threatening conditions that require close monitoring and support in order to maintain normal bodily functions [34].
3) **Hospitalized at transplant**: It represents whether the patient was already hospitalized when transplantation was done. In general, patients who are already in hospital tend to have lot of existing complications, infections, limited mobility, etc., which increases the risks of complications in addition to reducing the chance of successful outcomes.
4) **Lung Allocation Score at transplant**: The lung allocation score (LAS) is a numerical value used by UNOS to assign relative priority for distributing donated lungs for transplantation within the US. It takes into account various measures of a patient's health in order to direct donated organs towards the patients who would best benefit from a lung transplant [35]. This attribute represents the lung allocation score at the time of translplant.
5) **Lung Allocation Score at Listing**: Patients who are determined to be eligible for a lung transplant are placed on a waiting list. This waiting list is part of a national allocation system for donor organs run by the Organ Procurement and Transplantation Network (OPTN) [36]. This attribute represents the lung allocation score at the time of listing.
6) **Intubated at transplant**: Intubation refers to the insertion of a tube into an external or internal orifice of the body for the purpose of adding or removing fluids [37].
7) **GFR at transplant**: Glomerular filtration rate (GFR) is a test used to check how well the kidneys are working. Specifically, it estimates how much blood passes through the tiny filters in the kidneys (glomeruli) per minute [38]. This attribute represents the GFR at the time of transplant.
8) **Previous lung transplant**: It indicates whether the patient has undergone a lung transplant in the past.
9) **Age at transplant**: The age of the patient at the time of transplant.
10) **ECMO at transplant**: Extracorporeal membrane oxygenation (ECMO) is an extracorporeal technique of providing both cardiac and respiratory support oxygen to patients whose heart and lungs are so severely diseased or damaged that they can no longer serve their function [39]. This is maximal life support, but requires continuous infusion of heparin and blood circulating through large tubes that exit the body. This attribute represents the ECMO at the time of transplant.
11) **Previous lung transplant $<$90 days**: It indicates whether the patient has undergone a lung transplant within 90 days prior to the current transplant.
12) **Previous cardiac surgery**: It indicates whether the patient has undergone a cardiac surgical procedure in the past.

Using a predicted 1-year mortality risk $\geq$50% as a cutoff, the sensitivity of the final model for predicting 1-year mortality was 12%, the specificity 98%, and the F-measure 20%. Table I summarizes the performance of the final model comparing it to logistic regression, which is the most widely used technique in healthcare informatics for predictive modeling.

We believe that the preliminary results obtained in this work are quite encouraging and the fact that we can significantly reduce the number of attrbutes in the model without sacrificing accuracy motivates integration of such models in clinical decision making.

V. Conclusion and Future Work

In this workshop paper, we present our preliminary results of data mining on UNOS data on lung transplantation outcome. We evaluated nearly 50 classification schemes for predicting 1-year survival after the transplant. c-statistic of up to 0.68 was achieved. Further, feature selection techniques were able to significantly reduce the number of attributes in the model, incurring no cost in c-statistic. We believe that the resulting models can be very useful to aid physicians in decision making by providing them with patient-specific risk estimations.

Future work includes developing more sophisticated models for the studied outcome, and also exploring conditional outcome models using some post-transplant information (e.g. risk of 2-year mortality, given that the patient has already survived 1 year after transplant), and exploring the use of undersampling/oversampling to deal with unbalanced data. We also plan to do similar analysis for other transplants, and integrate the current and future work into healthcare and clinical decision making in practice, in the form of risk calculators.

VI. Acknowledgments

This work is supported in part by the following grants: NSF awards CCF-0833131, CNS-0830927, IIS-0905205, CCF-0938000, CCF-1029166, and OCI-1144061; DOE awards DE-FG02-08ER25848, DE-SC0001283, DE-SC0005309, DESC0005340, and DESC0007456; AFOSR award FA9550-12-1-0458.

References

[1] "Url: Lung transplantation, wikipedia, http://en.wikipedia.org/wiki/Lung_transplantation, accessed may 22, 2013."

[2] "Url: 2008 optn/srtr annual report, us scientific registry of transplant recipients, http://www.ustransplant.org/annual_reports/current/113_surv-new_dh.htm, accessed may 22, 2013."

Table I
ACCURACY OF FINAL MODEL AND LOGISTIC REGRESSION FOR PREDICTING 1-YEAR SURVIVAL AFTER LUNG TRANSPLANT

Evaluation Metric	Logistic Regression	Final Model
c-statistic (95% CI)	0.649 (0.630-0.668)	0.678 (0.660-0.696)
Sensitivity (Recall)	10.7%	12.3%
Specificity	98.2%	98.1%
Positive predictive value (Precision)	59.4%	61.0%
Negative predictive value	81.5%	81.8%
F-measure	18.2%	20.4%
Overall accuracy	80.7%	80.9%

[3] "Url: Lung transplantation, nlm nih, http://www.nlm.nih.gov/medlineplus/lungtransplantation.html, accessed may 22, 2013."

[4] "Url: 2008 u.s. organ and tissue transplant cost estimates and discussion, millman inc., http://publications.milliman.com/research/health-rr/pdfs/2008-us-organ-tisse-RR4-1-08.pdf, accessed may 13, 2013."

[5] "Url: Financing a transplant, tranplant living, http://www.transplantliving.org/before-the-transplant/financing-a-transplant/the-costs/, accessed may 13, 2013."

[6] J. Orens, M. Estenne, S. Arcasoy, J. Conte, P. Corris, J. Egan, T. Egan, S. Keshavjee, C. Knoop, R. Kotloff, F. Martinez, S. Nathan, S. Palmer, A. Patterson, L. Singer, G. Snell, S. Studer, J. Vachiery, and A. Glanville, "International guidelines for the selection of lung transplant candidates: 2006 update–a consensus report from the pulmonary scientific council of the international society for heart and lung transplantation," *J Heart Lung Transplant*, vol. 25, no. 7, pp. 745–55, 2006.

[7] "Url: United network for organ sharing, unos, http://www.unos.org/about/index.php, accessed may 13, 2013."

[8] V. N. Vapnik, "The nature of statistical learning theory," *Springer*, 1995.

[9] C. Bishop, *Neural Networks for Pattern Recognition*. Oxford: University Press, 1995.

[10] L. Fausett, *Fundamentals of Neural Networks*. New York, Prentice Hall, 1994.

[11] A. Pande, Y. Zeng, A. Das, P. Mohapatra, S. Miyamoto, E. Seto, E. Henricson, and J. Han, "Energy expenditure estimation with smartphone body sensors," in *8th International Conference on Body Area Networks (Bodynets) 2013*, 2013.

[12] R. E. Tarter, "Evaluation and treatment of adolescent substance abuse: A decision tree method," *The American journal of drug and alcohol abuse*, vol. 16, no. 1-2, pp. 1–46, 1990.

[13] S. Kim, K. Gopalakrishnan, and H. Ceylan, "Neural networks application in pavement infrastructure materials," in *Intelligent and Soft Computing in Infrastructure Systems Engineering*, 2009, pp. 47–66.

[14] B. Stenger, A. Thayananthan, P. H. Torr, and R. Cipolla, "Filtering using a tree-based estimator," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 2003, pp. 1063–1070.

[15] S. Jana, A. Pande, A. Chan, and P. Mohapatra, "Network characterization and perceptual evaluation of skype mobile videos," 2013.

[16] V. Omwando, A. Pande, Y. Zeng, and P. Mohapatra, "Evaluating perceptual video quality in 802.11n wlan with mobile clients," in *The 8th ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization (ACM WiNTECH) 2013*, 2013.

[17] R. Kohavi, "The power of decision tables," in *Proceedings of the 8th European Conference on Machine Learning*, ser. ECML '95, 1995, pp. 174–189.

[18] J. G. Cleary and L. E. Trigg, "K*: An instance-based learner using an entropic distance measure," in *In Proceedings of the 12th International Conference on Machine Learning*. Morgan Kaufmann, 1995, pp. 108–114.

[19] I. Witten and E. Frank, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann Pub, 2005.

[20] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[21] Y. Freund and L. Mason, "The alternating decision tree learning algorithm," in *Proceeding of the Sixteenth International Conference on Machine Learning*. Citeseer, 1999, pp. 124–133.

[22] H. George, "John and Pat Langley. Estimating continuous distributions in bayesian classifiers," in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, 1995, pp. 338–345.

[23] D. Hosmer and S. Lemeshow, *Applied Logistic Regression*. John Wiley and Sons, Inc., 1989.

[24] S. le Cessie and J. van Houwelingen, "Ridge estimators in logistic regression," *Applied Statistics*, vol. 41, no. 1, pp. 191–201, 1992.

[25] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," 1996.

[26] J. Friedman, T. Hastie, and R. Tibshirani, "Special invited paper. additive logistic regression: A statistical view of boosting," *Annals of statistics*, vol. 28, no. 2, pp. 337–374, 2000.

[27] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[28] T. Ho, “The random subspace method for constructing decision forests,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 832–844, 1998.

[29] J. Rodriguez, L. Kuncheva, and C. Alonso, “Rotation forest: A new classifier ensemble method,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 10, pp. 1619 –1630, oct. 2006.

[30] I. T. Jolliffe, *Principal Component Analysis*, 2nd ed. Springer, 2002.

[31] M. Hall, “Correlation-based feature selection for machine learning,” Ph.D. dissertation, Citeseer, 1999.

[32] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The weka data mining software: An update,” *SIGKDD Explorations*, vol. 11, no. 1, 2009.

[33] “Url: Functional status, npcrc, http://www.npcrc.org/resources/resources_show.htm?doc_id=376169, accessed may 22, 2013.”

[34] “Url: Intensive-care unit, wikipedia, https://en.wikipedia.org/wiki/Intensive-care_unit, accessed may 22, 2013.”

[35] “Url: Lung allocation score, wikipedia, http://en.wikipedia.org/wiki/Lung_allocation_score, accessed may 22, 2013.”

[36] “Url: What to expect before a heart transplant, national heart lung and blood institute, http://www.nhlbi.nih.gov/health//dci/Diseases/ht/ht_before.html, accessed may 15, 2013.”

[37] “Url: Intubation, wikipedia, http://en.wikipedia.org/wiki/Intubation, accessed may 15, 2013.”

[38] “Url: Glomerular filtration rate, medlineplus, http://www.nlm.nih.gov/medlineplus/ency/article/007305.htm, accessed may 15, 2013.”

[39] “Url: Extracorporeal membrane oxygenation, wikipedia, http://en.wikipedia.org/wiki/Extracorporeal_membrane_oxygenation, accessed may 15, 2013.”