VEGUETA Vo9 Issue 01,May 2021
ISSN NO: 1133-598X

Thoracic Surgery Outcome Prediction Made Easier

Dr. V Venkata Ramana[1,] B Swetha2, A Ram Praksh Reddy[3], Dr. V. Lokeswara Reddy[4]

[1,4] Professor, Department of [2,3] Asst. Professor,
Department of CSE, K.S.R.M College of Engineering(A), Kadapa

Abstract

Better monitoring of health outcomes may allow for enhancements to quality initiatives, healthcare management, and patient education. The Thoracic Surgery dataset includes patients who underwent extensive lung resections for primary lung cancer. When it comes to predicting survival in lung cancer patients following surgery, there is a lack of research and guidelines for practitioners interested in applying machine learning. Ability to rank and choose relevant characteristics is crucial when using machine learning for health outcome prediction. In this paper, we provide three attribute ranking and selection methodologies to improve the performance of algorithms in health outcomes research. By comparing the results of our proposed attribute ranking and selection methods to those of two scholarly publications, we demonstrate their efficacy.

Introduction

Modern physicians' efficiency and precision have been significantly boosted by the widespread use of computerized medical software. One such use case is measuring health-related outcomes. The importance of health outcomes in healthcare procurement and administration is undeniable. Cancer is becoming a leading killer in the majority of the world's nations. Lung cancer is now the leading indication for thoracic surgery[1]. Multiple methods were used by researchers to determine the kind of cancer before the onset of symptoms. Additionally, new approaches have been developed for the early prediction of the outcome of cancer treatment [2].As new medical technologies have emerged, huge cancer databases have been gathered and made accessible to scientists. However, the hardest part is making a correct diagnosis and prognosis of a condition. Thus, current research endeavors investigate the application of machine learning strategies for discovering and identifying models and relationships between them, from large datasets, the data is analyzed to extract useful information that supports disease augury, and to enhance models that predict patients' health with greater precision[3,4]. The effectiveness and efficiency of machine learning systems tend to deteriorate in the face of massive datasets. The computing time required to make predictions on high-dimensional data sets is greater. Complex datasets may be simplified via the use of attribute ranking and selection [5]. The field of machine learning has seen the presentation of several attribute and selection approaches. The primary goal of these techniques is to get rid of features that don't contribute to learning in any way, such as those that are deceptive, repetitive, or irrelevant. Selecting the most relevant qualities for class differentiation [6, 7] is the goal of attribute and ranking selection. Here's how the remainder of the paper is laid out: Sections 2 and 3 provide a brief overview of the ways in which machine learning algorithms and attribute ranking and selection techniques have been used to illness prognosis and prediction. In Section 4, we give the specifics of the data collection and the recommended approaches. Experimental findings are presented in Section 5. In Section 6, we talk about what we've learned.

Related Work

Thoracic surgery is the most extensive procedure for those with lung cancer. When deciding whether or not to operate, sawbones place a premium on the likelihood of survival. Considerations of short-term (e.g. post-operative complications, including death-rate in the first month) and long-term (e.g. survival for 1-5 years) risk and benefit for a patient is one of the common clinical decision challenges in thoracic surgery [8]. Disease prognosis and prediction have been used using a wide range of machine learning algorithms and attribute ranking and selection strategies during the last several decades. Machine learning methods for predicting cancer susceptibility, recurrence, and survival: a systematic literature review [2, 3]. Support Vector Machines, Bayesian Networks, Artificial Neural Networks, and Decision Trees are only few of the supervised machine learning approaches that K. Koura et al. [2] used to present prediction models with the goal of modeling cancer risk or patient outcomes.

Postoperative survival was predicted using boosted SVM in the study of Marcie Ziebach et al. [9]. The authors of the study addressed the issue of unbalanced data by using an oracle-based strategy to derive decision rules from the boosted SVM. Researchers Sandhog et al. [1] analyzed data from thoracic surgeries using six different classification methods (Naive Bays, J48, PART, One, Decision Stump, and Random Forest) and showed that Random Forest provides the greatest classification accuracy across all split percentages.

In another study [10], the authors examine the original and enhanced versions of four distinct machine learning methods (Nave Bayes, Simple logistic regression, Multilayer perception, and J48). Based on their findings, the boosted simple logistic regression strategy outperforms the other four machine learning methods studied by a little margin (84.53% to be exact).

Machine Learning

Machine learning is a branch of artificial intelligence which utilizes statistical, optimization and probabilistic techniques that allows computers to “learn” from past examples and to detect hard-to-discern patterns from large, noisy or complex data sets. These techniques have become a popular tool in medical diagnosis, which can find and identify models and relationships between them from large, noisy or complex datasets [3]. The inputs are the information about the patient’s age, gender, past medical history, past medical procedures, family medical history and current symptoms, while labels are the illnesses. In some cases, these inputs are missed because some tests haven’t been applied to the patient, so we do not apply machine learning techniques unless we confirm that the patient will give us valuable information. If the medical diagnosis is wrong, decision may lead to a wrong or no treatment, so machine learning is extremely used to diagnose and detect cancer [4].

More recently, it has been widely applied in the field of cancer prediction and prognosis which are differing from cancer detection and diagnosis. There are three types of cancer prediction and prognosis: One of them is prediction of cancer receptivity. In this type, one is trying to predict the probability of cancer progression before occurrence of the disease. Second type is the prediction of cancer recurrence by trying to predict the probability of redeveloping cancer after treatment and after a period of time during which the cancer cannot be detected. Third type is the prediction of cancer survivability by trying to predict an outcome which usually refers to life expectancy, survivability, progression and tumor-drug sensitivity.

These days, different types of cancer such as prostate, brain, cervical, esophageal, leukemia, head, neck, Breast and thoracic are appear to be compatible with machine learning prediction. The thoracic datasets is concerned with classification problem related to the post-operative life expectancy in the lung cancer patients [2-4]. In order to improve machine learning techniques when the datasets have a large number of features or attributes, attribute ranking and selection is used to identify the most relevant attributes and remove the redundant and irrelevant

Attributes from the dataset. Attribute ranking and selection algorithms can be divided into wrapper and filter methods. The wrapper methods select attributes based on an estimation of the accuracy according to target learning algorithm. After applying the learning algorithm, wrapper searches the feature space by removing some attributes and testing the effectiveness of attribute removing on the prediction metrics. The attribute which make important difference in learning process should be selected as high quality attribute, while filters methods estimate the quality of selected attributes independently from the learning algorithm. It depends on the statistical correlation between the set of attributes and the target attribute, since the value of correlation identify the importance of target attribute [6, 11]. By using filtering methods attributes can be ranked independently, and then according to the ranking result optimal subset of attributes can be selected [12].

The Proposed Method

Dataset description

Table 1 Thoracic surgery dataset attributes

ISSN NO: 1133-598X

Name	Description	Characteristics
DGN	Diagnosis - specific combination of ICD-10 codes for primary and secondary as well multiple tumors if any	Nominal
PRE4	Forced vital capacity - FVC	Numeric
PRE5	Volume that has been exhaled at the end of the first second of forced expiration - FEV1	Numeric
PRE6	Performance status - Zubrod scale	Nominal
PRE7	Pain before surgery	Binary
PRE8	Haemoptysis before surgery	Binary
PRE9	Dyspnoea before surgery	Binary
PRE10	Cough before surgery	Binary
PRE11	Weakness before surgery	Binary
PRE14	T in clinical TNM - size of the original tumor, from OC11 (smallest) to OC14 (largest)	Nominal
PRE17	Type 2 DM - diabetes mellitus	Binary
PRE19	MI up to 6 months	Binary
PRE25	PAD - peripheral arterial diseases	Binary
PRE30	Smoking	Binary
PRE32	Asthma	Binary
AGE	Age at surgery	Numeric
Risk1Y	1 year survival period - T value if died	Binary

Research Methodology

In this work, Version 3.7.12 of WEKA (Waikato Environment for Knowledge Analysis) toolkit [14] has been used for analysis. It is the product of the University of Waikato (New Zealand) and it is licensed under the GNU General Public License. WEKA is a popular suite of machine learning software written in Java, also it provides access to SQL database and process the result retrieved by a database query. We have run our experiments on a system with a 2.30 GHZ Intel(R) CoreTMi5 processor and 512 MB of RAM running Microsoft Windows 7 Professional (SP2).
Cross-validation (10 folds) has been used in this study to validate the results. In this model, the dataset is partitioned into complementary 10 equal sized subsets. The analysis is performed on 9 subsets (training) and validating the analysis on one subset (testing). Ten rounds of cross-validation are performed and in each round another subset 2 through 10 used as testing dataset. The validation results are averaged over the ten rounds in the final phase.

$$IG(Class, Attribute) = H(Class) - H(Class|Attribute) \quad (1)$$

Where H is the entropy which stands for the Greek Alphabet Eta. Symmetrical Uncertainty (SU) Attribute Evaluation [16] is used to evaluate the importance of an attribute by measuring the symmetrical uncertainty with Respect to the class. Symmetrical Uncertainty compensates for the inherent bias in Information Gain. Symmetrical Uncertainty is given by the following equation:

$$SU(Class, Attribute) = 2 * IG(Class, Attribute) / (H(Class) - H(Attribute)) \quad (2)$$

Relief-F (RF) attribute evaluation is used to rank the quality of features depending on how well their values differ from the cases that are close to each other. It is sensible to predict that a valuable feature should have different values between cases belong to different classes and have the same value for cases from the same class [17].
The aim of this paper is to analyze the effect of number of attributes on accuracy of machine learning techniques to solve the problem for prediction of the post-operative life expectancy in the lung cancer patients. Reducing the number of attributes and increasing the accuracy is required to minimize the computational time of prediction techniques.

Experimental Results

Performances of the methods were analyzed by using six metrics- accuracy, F measure, and ROC curve, Glean, TNR and TPR. Accuracy is the percentage of observations that were correctly predicted by the method. It was used to evaluate the performance of each algorithm.

$$\text{Accuracy} = TP + TN/(P + N) \quad (3)$$

Table 2 shows the confusion matrix which clarifies the prediction tendencies TP (True positive), TN (true negative), FP (false positive) and FN (false negative) of considered machine learning technique. Accuracy is not a reliable metric for the real performance of a machine learning technique, because it will yield misleading results if the data set is imbalanced (i.e. when the number of samples in different classes vary greatly). Since thoracic surgery data is imbalanced data with 70 true and 400 false instances we used F measure (F1 score), ROC curve, Glean, TNR and TPR. Where, F measure was used to test the accuracy depending on harmonic mean of precision & recall.

Table 2 Confusion matrix

		Predicted Outcome	
		P	N
Actual value	P	TP	FN
	N	FP	TN

$$\text{Fmeasure} = 2TP/(2TP + FP + FN) \quad (4)$$

While, ROC curve was also used as an effective method to evaluate the performance of predicted models by plotting the true positives against the false positives and area under the ROC curve is used for predicting accuracy of models. The Glean (geometric mean) is a widely used quality rate and is defined:

$$Gmean = \sqrt{TPR \times TNR} \quad (5)$$

$$TNR = TN/(TN + FP) \quad (6)$$

$$TPR = TP/(TP + FN) \quad (7)$$

Table 3 shows the accuracy of Na¨ıve Bays, Simple Logistic Regression, and J48 and Multilayer Perception techniques with and without using attribute ranking methods. Also, it shows the accuracy of boosted Na¨ıve Bays, boosted Simple Logistic Regression, boosted J48, boosted Multilayer Perception for prediction of post-operative life expectancy after Thoracic Surgery. Results show that using IG and SU as ranking methods before applying Na¨ıve Bays gives the better accuracy than applying Na¨ıve Bays without using ranking methods and with boosted Na¨ıve Bays. Also, in the case of applying Simple logistic, using the three ranking methods gives better accuracy than applying Simple logistic without using ranking methods and with boosted Simple logistic.

Table 3 Prediction accuracy comparison of machine learning techniques using thoracic surgery data set

ML techniques	Method	Accuracy
Naive Bayes	Original	77.74
	Boosted	78.32
	SU	82.12
	RF	77.74
	IG	82.13
Simple logistic	Original	84.55
	Boosted	84.53
	SU	84.68
	RF	84.68
	IG	84.68
Multilayer Perceptron	Original	80.91
	Boosted	80.7
	SU	81.27
	RF	81.28
	IG	81.28
J48	Original	84.64
	Boosted	79.34
	SU	84.46
	RF	84.47
	IG	84.47

Table 4 shows the F measure and ROC curve of Na¨ıve Bays, Simple Logistic Regression, J48 and Multilayer Perception techniques with and without using attribute ranking methods. Also, it shows the F measure, ROC curve of boosted Na¨ıve Bays, boosted Simple Logistic Regression, boosted J48, boosted Multilayer Perception for prediction of post-operative life expectancy after Thoracic Surgery. Results show that applying Na¨ıve Bays without using ranking methods gives the better F measure than using the three ranking methods before applying Na¨ıve Bays and with boosted Na¨ıve Bays, but, boosted Na¨ıve Bays gives the best ROC curve. In the case of applying Simple logistic, it gives the same results for the F measure in all methods, but using the three ranking methods gives the best ROC curve. In the case of applying Multilayer Perception, using SU and IG ranking methods gives the best F measure and the best ROC curve. In the case of applying J48, boosted J48 gives the best F measure but applying J48 with and without using ranking methods gives better ROC curve than boosted J48.

Table 4 Prediction measures comparison of machine learning techniques using thoracic surgery data set

ML techniques	Method	F measure	ROC
Naïve Bayes	RF	0.06	0.66
	SU	0.06	0.66
	IG	0.07	0.67
	Boosted	0.12	0.6
	Original	0.13	0.68
Simple logistic	RF	0	0.5
	SU	0	0.5
	IG	0	0.5
	Boosted	0	0.61
	Original	0	0.53
Multilayer Perceptron	RF	0.2	0.58
	SU	0.24	0.55
	IG	0.24	0.55
	Boosted	0.18	0.56
	Original	0.22	0.6
J48	RF	0.02	0.5
	SU	0	0.5
	IG	0	0.51
	Boosted	0.18	0.61
	Original	0	0.5

Table 5 Performance evaluation of boosted SVM vs. SVM

Method	TPR	TNR	Gmean
Boosted SVM(BSI)	60	72	65.73
SVM (IG)	44.3	99.8	66.49
SVM (SU)	44.3	99.8	66.49
SVM (RF)	51.4	99.8	71.62

Table 5 shows the TPR, TNR, and Glean for support vector machine after applying the three attribute ranking and selection methods and boosted support vector machine. The RF gives the best prediction quality where it has the higher Glean value. It shows also that the proposed methods give better Glean and TNR than Boosted SVM but Boosted SVM gives better TPR.

Conclusion

As a means of better predicting survival rates for patients with lung cancer after thoracic surgery, this research compares three attribute ranking and selection strategies. We have compared five machine learning approaches before and after using attribute ranking and selection methods. Results demonstrate that attribute ranking and selection can outperform boosting in increasing prediction accuracy. Future work may include the use of additional attribute selection and machine learning approaches in order to improve the prediction model performance of the dataset...

References

This is according to [1] V. Sandhog, S. A. S. Piranha, S. Vein, and M. Hemalatha. International Journal of Computer Technology & Applications, Vol. 5, Pp. 578–586 (2014), "Thoracic surgery analysis using data mining techniques."

Computational and Structural Biotechnology Journal, Volume 13, Pages 8–17, 2015, "Machine Learning Applications in Cancer Prognosis and Prediction" by KonstantinaKourou, Themis P. Exarchos, Konstantinos P. Exarchos, Michalis V. Karamouzis, and Dimitrios I. Fotiadisa.

[3] KwetisheJoro Danjuma. (2015), "Performance evaluation of machine learning algorithms in post-operative life expectancy in the lung cancer patients", IJCSI International Journal of Computer Science Issues, Vol. 12, No. 2, pp. 145-157.

pp.189-199 .

In 2006, Joseph A. Cruz and David S. Wishart published "Applications of machine learning in cancer prediction and prognosis" in Cancer Informatics, Volume 2, Pages 59–77.

"A hybrid feature selection method to improve performance of a group of classification algorithms," by Mehdi Naseriparsa, Amir-Masoud Bidgoli, and Touraj Varaee (2013), International Journal of Computer Applications, Volume 5.

69, No. 17,pp.28-35.

[6] Pinar Yildirim. (2015), "Filter based feature selection methods for prediction of risks in hepatitis disease", International Journal of Machine Learning and Computing, Volume 5, Issue 4, Pages 258-263.

Science and Information Conference (SAI), Science and Information Conference (SAI), pp.372-378, 2014, "A survey of feature selection and feature extraction techniques in machine learning" [7].

[8] Koklu, Murat, HumarKahramanli, and NovruzAllahverdi. (2015), "Applications of rule based classification techniques for thoracic surgery", TIIM international conference, 27-29 May 2015, Italy, pages 1991-1998.

"Boosted SVM for extracting rules from imbalanced data in application to prediction of the post-operative life expectancy in the lung cancer patients," M. Zi,eba, J. M. Tomczak, M. Lubicz, and J. 'Swi,atek. (2014), Applied Soft Computing, Volume 14, Pages 99-108.

[10] "Predicting outcome of thoracic surgery by data mining techniques",IJARCSSE, Vol. 5, No. 1,pp. 7-10, Md. AhasanUddinHarun and Md. NureAlam., 2015.