Application of Classification Methods to Elective Surgical Cases Cancellation Detection

LI Feng[1, a *], Li Luo[1, b] Renrong Gong[2]

[1]Business School of Sichuan University, Chengdu, China

[2]West China Hospital of Sichuan University, Chengdu, China

[a]523150676@qq.com, [b]luolicc@scu.edu.cn

Keywords: Surgery Cancellation, Operating Room Management, Decision Tree, Bayes Network, Classification Techniques

Abstract. The case cancellation in the operating room can cause multi-faceted troubles, so it is difficult for the operating room manager to detect potential cancelled cases. The objective of this study is to build classification models like Decision Tree and Bayes Network to assist the operating room manager to detect the potential cancelled cases. After data acquisition and data preprocessing, classification models are trained. As a result, Decision Tree and Bayes Network outperform SVM and Neural Networks in terms of classification accuracy.

Introduction

Cancelled surgeries are inconvenient to patients, physicians, and staff, leading to decreased patient satisfaction and diminished staff morale [1]. For the purpose of solving the troubles of hospitals and patients caused by cancellation, a great many previously published papers [1,2,3] have dedicated to analyzing the reasons for cancellation and putting forward suggestions to minimize the number of cancellation. In practical work, detecting high-risk surgery cancellation is complicated but crucial, which demands more automatic classification methods and techniques that are capable of detecting high-risk surgery cancellation from large databases.

Nowadays, classification techniques have been widely used in a variety of fields, including fraud detection, customer relationship management and human resource management, etc. Our aims at carrying this study is to design, develop and test classification models in order to detect potential cancelled surgeries.

This paper is organized as follows: presented in the next station is a detail description of data sources and data preprocessing, followed by an introduction to the methods we adopt and the modeling steps. Then, the paper puts forth the classification results and compares the performance of classifiers. The last section concludes the paper.

Data Sources and Data Preprocessing

After acquiring the admission of West China Hospital of Sichuan University, we collect cancellation data from urology department over 12-month period from January 2013 to January 2014. There are two datasets that provide information about patient demographic, surgery needs submitted by surgeons and surgeries that are ultimately performed as planned. We combine these datasets by using the same patient ID and surgery day. The resultant dataset has 5125 integrated elective surgery data.

There are 24 variables in each record, including patient ID, operative types and surgeon, etc. Actually, operative types have a profound influence on surgery cancellation, but we abandon this variable for the following two reasons: first, there are about 500 types in 5125 surgeries which are mainly hybrid surgeries, and too careful division may generate very small sample in each type, which is harmful to the construction of classification models, at the same time, the representation of the general law reduced greatly; second, in this hospital, disease is not the main reason for cause cancellation, the objective of our research is to detect the cancellation which is brought out by some administrative reasons, thus excluding the disease factor from consideration. Then, chi-square test [4] and one-way

analysis of variance are employed to find out which variables are strongly associated with the target variable. As a result, such attributes as “Surgeons”, “Sequence of surgery”, “Days after admission”, “Surgery room”, “Cancelled before” have strong association with the target variable, hence, making the number of attributes reduce from twenty four to five, accordingly. The variables detail are presented in Table 1.

The correct size of the training set is an important parameter in classification experiment [5]. To avoid the occurrence of overfitting, we sample training set and test set in the proportion of 60:40, with the partition of each set determined by the random number generator.

Table 1 Adopted Variable Detail

Variable name	Variable type	Variable detail	Variable meaning
Surgeon	Discrete	14 categorical values	The doctor in charge of the surgery.
Sequence of surgery	Range	[1-10]	The sequence of surgery.
Days after admission	Range	[0-90]	Number of days after admission.
Surgery room	Set	[1,2,…,7]	The rooms where surgeries are performed.
Cancelled before	Flag	Yes, no	The surgery has been cancelled before.
Cancellation	Flag	Yes, no	Whether the surgery be performed finally.

Methods

Commonly used classification methods are Decision Tree (DT), Bayes network (BN), Neural Network (NN) and Support Vector Machines (SVM), etc. Decision Tree method has the advantage that it can produce an easily interpretable top-down tree with high accuracy of classification, and the most important variable will float to the top of tree, while techniques like NN can’t interpret its result. Considering that one of our target is to help the hospital manager have a better understanding of the principle of the classification method, we take Decision Tree, but not exclusively, as the preferred method, and make a comparison with other classification methods.

Decision Tree, which can be binary or multi-way, is composed of a root node, internal nodes, leaf nodes, and branches. The modeling process can be divided into the following steps: (1) Creating node N; (2) If the samples are all in the same class, the procedure is over with N as a leaf node and labelled as the most common category; (3) If the samples are in the different classes, select a test attribute which can classify samples by various decision tree algorithm in order for the biggest reduction of diversity; (4) Creating a branch for every value of the test attribute and a decision tree recursively, the recursive procedure is over when all the samples in the given node are in the same class or there is no remaining properties that can be used for partition. Decision Tree algorithms rely on various ways of measurement of subset’s impurity, including information gain, gain ratio and distance measure, etc. Choosing the appropriate way of measurement has a great influence on results. Considering the fact that the algorithm of CART [6] can only produce a binary tree, and excessive merge will give rises to logical chaos when compared with the practical situation of cancellation, we abandon this method and take CHAID as our final choice.

Rather than with surgeries which are performed as planned, we care more about the potential cancelled cases, so we utilize misclassification costs and set the cost of “abandon true” 5 times as high as that of “fetch wrong”. To ensure that the calculated error rate comes close to the true error rate, cross-validation is employed and the number of folds is 10[7]. A fully grown Decision Tree does not have the best ability of classification due to the fact that a fully grown tree’s description of the training set is too precise. With the growth of decision tree, the representativeness of the whole dataset’s law is decreasing continuously. We take as a solution the post-pruning method, which, on the basis that Decision Tree is fully grown, prunes those sub trees that don’t have general representativeness according to some rules. Many attempts have been made to get the best result with a final pruning severity of 0.95.

Besides, SVM, NN and BN are also adopted in order to find out the best prediction model. Support Vector Machines is developed based on the structure risk minimization principle and VC dimension theory which can handle non-linear decision boundaries of arbitrary complexity[8]; Neural Networks are nothing more than function approximation tools which learn the relationship between independent variables and dependent variables, much like regression or other more traditional approaches[9]; On the basis of prior probability, Bayes Network makes the most of the discrimination function to adjust the prior probability and get the probability estimation on the classification of samples. Due to the fact that every model has various demands on modeling data, for example, the Neural Networks can only deal with numerical data, so we preprocess the dataset in accordance with the corresponding data requirements. In addition, surgeries performed as planned ultimately take most of the proportion (85%), which means the building model will tend to judge the coming surgery as a cancelled case, so the imbalance of samples is adverse to the models. Adopting the over-sampling solution, we extract all the cancelled surgeries and a small part of the performed surgeries to make each sample achieve a state of balance. Consequently, 812 cancelled cases and 825 performed cases are extracted from the whole dataset when we build models NN, BN and SVM.

Results

Import the whole dataset to SPSS Clementine 12.0 and create classification models according to the aforementioned methods. Finally, the size of the tree is fifteen, which contains one root node, five internal nodes and nine leaves. The observation of decision tree provides clues for the construction of some rules for the distinction between cancelled cases and surgeries performed as planned. For example, one of the rules says that if the number of days after admission is within two and the surgeon's number is eight, the surgery is a potential cancelled case. Similar rules can be derived from the Decision Tree (Figure 1).

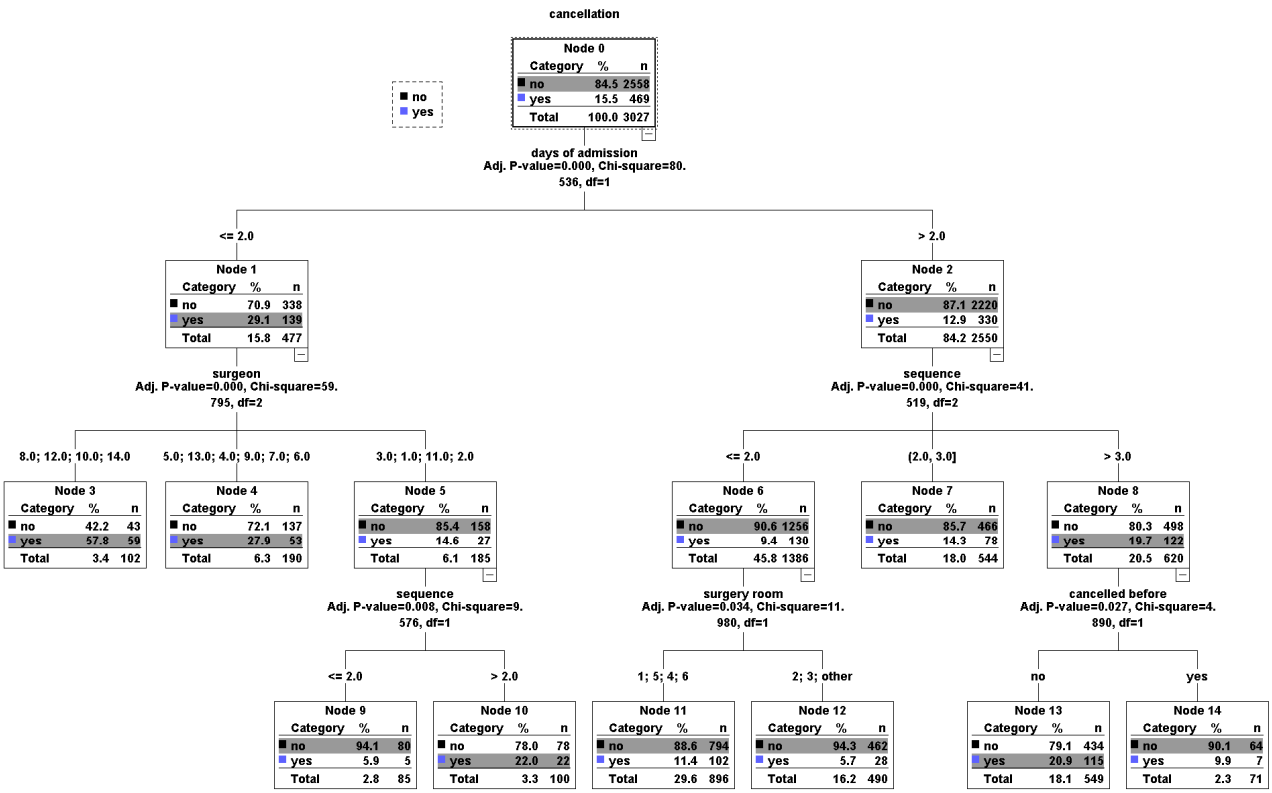

Figure 1 Decision Tree Based on CHAID

Judging from the final Decision Tree, four variables are finally chosen, while "Anesthesia" is casted away. Furthermore, we can notice that "Days after admission" is chosen as the first attribute to divide

the root node, which means “Days after admission” is judged as the most important variable that is strongly related with the target variable. According to gain chart (Figure 2), in which the gain curve rises steeply at the beginning and becomes flat gradually, the decision tree model shows good stability both on training set and test set.

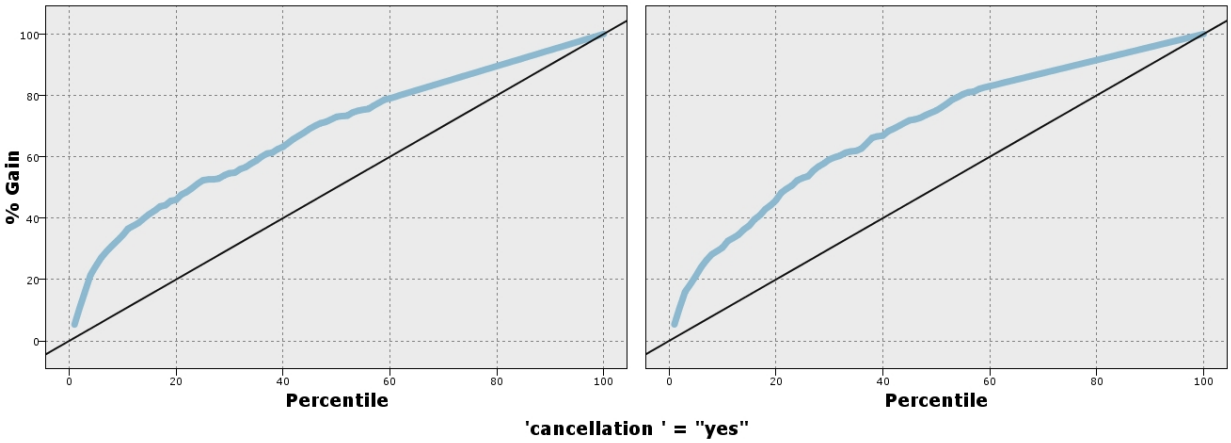

Figure 2 Gain Chart

From the confusion matrix, it is not difficult to find that the Bayes Network and Decision Tree have a high accuracy on target category, while Neural Network and SVM produce a much higher classification accuracy in the class “No” than Decision Tree and Bayes Network. Taking into account that the class “Yes” is what we really care about, we can draw the conclusion that Bayes Network and Decision Tree outperform Neural Network and SVM in terms of the detection of potential cancelled cases.

Decision Tree is built not only for the purpose of detecting the potential cancelled cases, but also to help the hospital manager acquire a better knowledge of cancellation and thus take some feasible measures. For example, strict preoperative examination should be performed on patients for whom the number of days after admission is less than two, to improve the efficiency of the operations in surgery rooms.

Table 2 Confusion Matrics for Four Classification Models

Decision Tree	Predicted				Neural Network	Predicted			
Actual	Yes	No	Total	Corr Rate	Actual	Yes	No	Total	Corr Rate
Yes	494	318	812	60.8%	Yes	48	764	812	5.9%
No	1239	3074	4313	71.3%	No	7	818	825	99.2%
Total	1733	3392	5125	69.6%	Total	55	1582	1637	52.9%
Bayes Network	Predicted				SVM	Predicted			
Actual	Yes	No	Total	Corr Rate	Actual	Yes	No	Total	Corr Rate
Yes	484	328	812	59.6%	Yes	14	798	812	1.7%
No	325	500	825	60.6%	No	20	805	825	97.6%
Total	809	828	1637	60.1%	Total	34	1603	1637	50.0%

Conclusion

The data mining on cancellation of surgery, which can play a supporting role for the operating room manager, can dig out the knowledge hidden in the vast data of surgeries. We apply classification techniques like Decision Tree, Bayes Net, Neural Network and SVM to the forecasting of the cancellation and to providing a new method for the manager. The result shows that the accuracy of Decision Tree and Bayes Network are apparently higher than those of SVM and Neural Network. There are three main limits of our research: first, although the built models are competent at the classification job, but the accuracy need to be improved; second, with the models established for specific data and the results may only be applied to the specific hospital, the built models may not be applicable to other hospitals; third, in order to build a more comprehensive model, not only the administration factor, but also the disease factor should be taken into consideration, with the practical situation being more complicated than the supposed one. For the fact that every single model has its bias on the classification, we should build some ensemble models to avoid the bias and improve the forecasting accuracy.

Acknowledgment

This study is sponsored by the Nature Science Foundation of China (71532007, 71131006, 71172197) and Central University Fund of Sichuan University Number skgt201202. Here we acknowledge the support from West China Hospital of Sichuan University.

References

[1] A. R. Tait, T. VoepelLewis, H. M. Munro, H. B. Gutstein, and P. I. Reynolds, Cancellation of Pediatric Outpatient Surgery: Economic and Emotional Implications for Patients and Their Families, Journal of Clinical Anesthesia, 9 (1997), 213-19.

[2] Joshua L. Argo, Catherine C. Vick, Laura A. Graham, Kamal M. F. Itani, Michael J. Bishop, and Mary T. Hawn, Elective Surgical Case Cancellation in the Veterans Health Administration System: Identifying Areas for Improvement, American Journal of Surgery, 198 (2009), 600-06.

[3] John Geriaine, Six Sigma Plan Delivers Stellar Results, Materials management in health care, 16 (2007), 20-26.

[4] H. Liu, and R. Setiono, Chi2: Feature Selection and Discretization of Numeric Attributes, in Seventh International Conference on Tools with Artificial Intelligence, Proceedings, ed. by J. F. Vassilopoulos, 1995), pp. 388-91.

[5] F. Bonchi, F. Giannotti, G. Mainetto, and D. Pedreschi, Using Data Mining Techniques in Fiscal Fraud Detection. ed. by M. Mohania and A. M. Tjoa, Data Warehousing and Knowledge Discovery. First International Conference, Dawak'99. Proceedings, 1999), pp. 369-76.

[6] C. H. Yeh, Classification and Regression Trees (Cart), Chemometrics and Intelligent Laboratory Systems, 12 (1991), 95-96.

[7] R. Kohavi, A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection, IJCAI-95. Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (1995), 1137-43 vol.2.

[8] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, Gene Selection for Cancer Classification Using Support Vector Machines, Machine Learning, 46 (2002), 389-422.

[9] K. A. Smith, and J. N. D. Gupta, Neural Networks in Business: Techniques and Applications for the Operations Researcher, Computers & Operations Research, 27 (2000), 1023-44.

