Performance Evaluation of Supervised Machine Learning Classifiers for Predicting Healthcare Operational Decisions

Technical Report · January 2017

DOI: 10.13140/RG.2.2.26371.25127

2 authors:

Muhammad Zain Amin
University of Engineering and Technology, Lahore

Amir Ali
Warsaw University of Technology

Performance Evaluation of Supervised Machine Learning Classifiers for Predicting Healthcare Operational Decisions

Muhammad Zain Amin
Co-Founder of Wavy AI
Mountain View, California

Amir Ali
Co-Founder of Wavy AI
Mountain View, California

Abstract. This paper describes a healthcare operational decision making system based on machine learning classifiers to predict the decisions in comparison to the actual decisions made by the doctor during the healthcare operations. Most of the supervised machine learning classification and optimization techniques is utilized in this type of decision making system. This system can help the doctor make the best decisions. We testify this system on caesarian section which is the most commonly performed obstetric operation in the world to help saves mother and baby. This system helps us to predict when we should use surgery. This study explains utilization of machine learning algorithms in determination of medical operation methods. The results show that k nearest neighbors and Random Forest for this case study generates accuracy of 95.00 % respectively.

Keywords: Machine Learning classifiers, Healthcare Decision Making, K nearest neighbor, Random Forest

1. INTRODUCTION

Today's healthcare needs effective methods and research methodologies to save lives, reduce the cost of the healthcare services and early discoveries of contagious diseases. Machine learning techniques can enable healthcare organizations to predict trends in patient conditions and their behaviors. Recent findings in healthcare sector led to the collection of large size of rich data. McKinsey estimates that big data and machine learning could generate a value of $100 billion annually based on better decision making, optimized innovation and improved efficiency of clinical trials. Extracting useful knowledge and regularities from datasets can provide a major opportunity for practical use to improve healthcare. Knowledge acquired in this manner can be used to predict trends of patient's condition in shortest possible time and reduce the cost of healthcare services.

Information technologies are being increasingly implemented in healthcare organizations in order to respond to the needs of doctors in their operational decision making activities. Machine Learning can not only help in decision making in emergency medical situations but also in general primary care. Moreover machine learning techniques can also be used to help physicians diagnose patients especially in cases when outcomes are hardly to predict and choose the best operation method [1].

The term machine learning was introduced in 1959 by Arthur Samuel. The above points indicate that there is a great need of new computational theories and tools to extract information from large volume of datasets [2]. The role of machine learning within the field of data mining and processing of large datasets increased with the discoveries of several algorithms such as support vector machine methods in 1990s [3].

The first use of machine learning and data analytics in healthcare information systems was satisfied since 1970s [4]. Machine learning control limitations of people such as subjectivity due to fatigue and provide indications for the decision making process [5].

The organization of this paper is as follow. Section 2, describes machine learning in healthcare. Section 3 includes description of supervised machine learning classifiers such as k nearest neighbors, random forest, logistic regression, naive bayes, support vector machine. Section 4 describes the caesarian section followed by methodology as section number 4. Section 5 explains the results and discussion section.

2. MACHINE LEARNING IN HEALTHCARE

Machine learning in today's healthcare is unavoidable. Optimists predict that machine learning and artificial intelligence will diagnose disease better and earlier, treat illness more precisely and engage patients more efficiently in future healthcare. Recent advancements in machine learning have demonstrated that machine learning can create algorithms that perform on par with human physicians.

In recent years instances in healthcare such as medical image processing and analyzing, predicting healthcare operational decisions, dosage trials for intravenous tumor treatment and detection and management of prostate cancer.

3. SUPERVISED MACHINE LEARNING CLASSIFIERS

Machine learning focuses on the development of computational algorithms that can access data in order to look for patterns in data and make better predictions about future instances based on the instances that we provide without being explicitly programmed. The machine learning is called supervised if known labels are given with instances in the training phase, whereas instances are unlabeled in unsupervised machine learning. This section focuses on machine learning techniques such as support vector machine (SVM), naïve Bayes classifier (NB), random forest (RF), k-Nearest neighbor (KNN) and logistic regression.

3.1 RANDOM FOREST (RF)

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. Random decision forests correct for decision trees' habit of over fitting to their training set.

3.2 LOGISTIC REGRESSION (LR)

It is a statistical method for analyzing a data set in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible outcomes). The goal of logistic regression is to find the best fitting model to describe the relationship between the dichotomous characteristic of interest (dependent variable = response or outcome variable) and a set of independent (predictor or explanatory) variables.

3.3 NAIVE BAYES (NB)

It is a classification technique based on Bayesian network with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability. NB is based on probability estimations, called a posterior probability.

3.4 K-NEAREST NEIGHBOURS (kNN)

The k-nearest-neighbors algorithm is a classification algorithm, and it is supervised: it takes a bunch of labelled points and uses them to learn how to label other points. To label a new point, it looks at the labelled points closest to that new point (those are its nearest neighbors), and has those neighbors vote, so whichever label the most of the neighbors have is the label for the new point (the "k" is the number of neighbors it checks).

3.5 SUPPORT VECTOR MACHINE (SVM)

A Support Vector Machine (SVM) is a discriminative classifier which can be used for both classification and regression problems. The goal of SVM is to identify an optimal separating hyperplane which maximizes the margin between different classes of the training data. In other words, given labeled training data (*supervised learning*), the algorithm outputs an optimal hyperplane which categorizes new examples to create the largest possible distance to reduce an upper bound. Supports Vectors are simply the co-ordinates of data points which are nearest to the optimal separating hyperplane provide the most useful information for SVM classification. In addition, an appropriate kernel function is used to transform the data into a high-dimension to use linear discriminate functions.

4. CAESARIAN SECTION

Caesarian section is a widely performed obstetric surgery to deliver a baby. Caesarean rates in the U.S. have risen considerably since 1996 [6]. The procedure increased 60% from 1996 to 2009. In 2017, the Caesarean delivery rate was 32% of all births [7]. Nearly one in three pregnant women in the United States gave birth by Caesarean section in 2015, according to the most recent birth statistics from the Centers for Disease Control and Prevention (CDC). C-sections can help women who are at risk for complications avoid dangerous delivery-room situations and can be a lifesaver in an emergency. The World Health Organization officially withdrew its previous recommendation of a 15% C-section rate in June 2010. Their official statement read, "There is no empirical evidence for an optimum percentage. What matters most is that all women who need caesarean sections receive them" [8].

UK National Confidential Enquiry have been purposed four categories of exigency (Table 1) into Patient Outcome and Death (NCEPOD), and backed by the UK Royal College of Obstetricians and Gynecologists (RCOG) and the UK Royal College of Anesthetists (RCA) [9].

Category	***Criterion***
1	Immediate threat to the life of the woman or fetus
2	Maternal or fetal compromise that is not immediately life threatening
3	No maternal or fetal compromise but early delivery required
4	Delivery timed to suit woman and staff (elective)

Table 1.Categories of emergency for caesarian section

5. METHODOLOGY

Fig. illustrates the methodology of decision making system based on supervised machine learning classifiers.

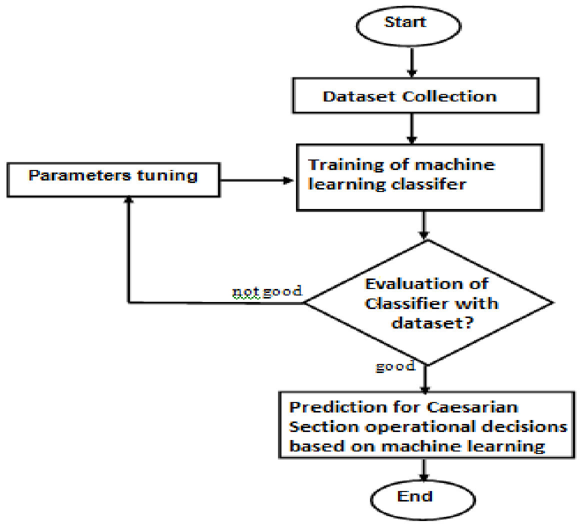

Fig.1, A methodology of decision making system based on machine learning classifiers

As shown in Fig 1, three steps are required to be accomplished in order to predict decisions: dataset collection, training of machine learning classifiers and evaluation of machine learning classifiers.

5.1 DATASET COLLECTION

The dataset of caesarian section was collected from the "Application of Decision Tree Algorithm for Data Mining in Healthcare Operations: A Case Study" [10]. The dataset consists of 80 instances, composed of five attributes which are the most important characteristics of delivery problems namely age, number of pregnant, delivery time, blood pressure and heart status. These attributes are classified in Table.

Table 2. ATTRIBUTES CLASSIFICATION TABLE

Attributes	**Classification**
Age	Numerical
No. of Pregnant	Numerical
Delivery Time	Premature, Timely, Latecomer
Blood Pressure	Low, Normal, High
Heart Status	Apt, Inept
Caesarian	No, Yes

5.2 TRAINING OF MACHINE LEARNING CLASSIFIERS

The training of supervised machine learning classifiers is known as classification. Training is based on the task of concluding a classifier from a labelled caesarian section dataset. The supervised machine learning algorithm examines the training caesarian section dataset and produces a classifier which can effectively predict the decisions for the caesarian section.

5.3 EVALUATION OF CLASSIFIERS:

During the data training process, the machine learning algorithm generates the classification model by using the training data. In this study training dataset was used to evaluate the machine learning classifiers used for the prediction of the C-section surgery.

Table 3. CONFUSION MATRIX

	Predictive Positive	**Predictive Negative**
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

Important measures were extracted from the confusion matrix in order to accurately evaluate the machine learning classifiers. In addition to correct classification rate or accuracy other measures that were used to evaluate the machine learning classifiers are True Positive Rate (TPR), False Positive Rate (FPR), Precision, Recall, F1 score and ROC area.

Table 4. THE MEASURES USED FOR EVALUATING MACHINE LEARNING CLASSIFIERS

Measure Name	**Formula**
Correct Classification Rate	$= \frac{TP + TN}{TP + FP + FN + TN} (\%)$
True Positive Rate	$= \frac{TP}{TP + FN}$
False Positive Rate	$= \frac{FP}{TN + FP}$
Precision	$= \frac{TP}{TP + FP}$
Recall	$= \frac{TP}{TP + FN}$

6. RESULTS AND DISCUSSION

The Caesarian Section Dataset was obtained from "Application of Decision Tree Algorithm for Data Mining in Healthcare Operations: A Case Study" [10] to evaluate the machine learning classifiers. In the caesarian section dataset 80 instances were gathered and used for evaluating and training the supervised machine learning classifiers used for predicting healthcare operational decisions. In our training dataset, instances with 'yes' decisions for caesarian section surgery is assigned to a class with label 1 and instances with 'no' decisions assigned to a class with label 0.

The performances in terms of correct classification rate (CCR), True Positive Rate (TPR), False Positive Rate (FPR), Precision, Recall, F1 score and ROC area of logistic regression, naïve bayes, K nearest neighbors, random forest and support vector machine were compared together and discussed below.

In this research, performances of supervised machine learning classifiers were calculated using WEKA software.

Table 5. ACCURACY OF COMPUTING ON DATASET

Machine Learning Classifiers	TP Rate	FP Rate	Precision	F1 Rate	MCC Rate	ROC Area
SVM	0.763	0.252	0.762	0.762	0.512	0.755
RF	0.950	0.052	0.950	0.950	0.898	0.994
NB	0.763	0.229	0.771	0.764	0.527	0.842
kNN	0.950	0.037	0.955	0.950	0.904	0.995
LR	0.775	0.235	0.775	0.775	0.540	0.875

We further calculated the performance of classifiers in terms of kappa static, mean absolute error, root mean squared error, relative absolute error and root relative squared error.

Table 6. EVALUATION RESULT ON TRAINING DATA

	SVM	RF	NB	LR	kNN
Correctly Classified Instances	61	76	61	62	76
Incorrectly Classified Instances	19	04	19	18	04
Kappa Statistic	0.5122	0.8977	0.5232	0.5396	0.8992
Mean Absolute Error	0.2375	0.1879	0.3292	0.2856	0.0604
Root Mean Squared Error	0.4873	0.2346	0.4012	0.3791	0.1585
Relative Absolute Error	48.56	38.42	67.31	58.40	12.34
Root Relative Squared Error	98.58	47.45	81.52	76.69	32.06
Total Number of Instances	80	80	80	80	80

As you see, both the k nearest neighbors and Random Forest predicted correct result for 95% cases. So the information gained with the use of machine learning algorithms can be useful to make successful decisions that will improve success of healthcare decision making and health of the patients.

7. CONCLUSION

Machine learning has great importance for area of healthcare, and it represents comprehensive process that demands thorough understanding of needs of the healthcare organization regarding operational decision making. In this study, the machine learning techniques was used for selecting the most significant features to be utilized in predicting caesarian section accurately. According to our paper, machine learning algorithms namely RF, kNN, SVM, LR, and NB were applied to calculate the Performance Evaluation of Supervised Machine Learning Classifiers for predicting caesarian section operational decisions. The experiment results show that both RF and kNN achieved the best accuracy rates by predicting 95 cases correctly. The success of random forest and k nearest neighbors is measured by its accuracy and in future we can increase the reliability of caesarian section decision making by further increasing the number of instances and adding some more core attributes.

8. REFERENCES

[1] Boris Milovic, Milan Milovic. Prediction and Decision Making in Health Care using Data Mining, International Journal of Public Health Science (IJPHS), Vol. 1, No. 2, December 2012, pp. 69~78

[2] Fayyad, U., Shapiro, G. P., & Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. American Association for Artificial Intelligence, 37-54.

[3] Kantardzic, Mehmed. Data Mining: Concepts, Models, Methods, and Algorithms. John Wiley & Sons, 2003.

[4] H. Jiawei and K. Micheline, Data Mining: Concepts and Techniques, vol. 2, Morgan Kaufmann Publishers, 2006.

[5] Candelieri, A., Dolce, G., Riganello, F., & Sannita, W. G. (2011). Data Mining in Neurology. In KnowledgeOriented Applications in Data Mining (pp. 261-276). InTech.

[6] Brady E. Hamilton, Ph.D.; Joyce A. Martin, M.P.H.; and Stephanie J. Ventura, M.A., Division of Vital Statistics, Births: Preliminary Data for 2007, National Vital Statistics Report.

[7] Births: Provisional Data for 2017 *USA. CDC. May 2018*. Retrieved 18 May 2018.

[8] World Health Organization (WHO) statement “Should there be a limit on Caesareans?". BBC News. 30 June 2010.

[9] Andrew Simm, Darly Mathew, Caesarian section: techniques and complications, Obstetrics, Gynaecology & Reproductive Medicine, Volume 18, Issue 4, April 2008, Pages 93-98.

[10] Farhad Soleimanian Gharehchopogh, Peyman Mohammadi, Parvin Hakimi, Application of Decision Tree Algorithm for Data Mining in Healthcare Operations: A Case Study, International Journal of Computer Applications (0975 – 8887) Volume 52 – No. 6, August 2012,Pages 21-26.