ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and CommunicationEngineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

Prognosis of Failed Back Surgery Syndrome Based On Feature Extraction Method

Alda Kika[1], Ridvan Alimehmeti[2]

Department of Informatics, Faculty of Natural Science, University of Tirana, Tirana, Albania[1]

Neurosurgical Department at University of Medicine, Tirana, Albania[2]

ABSTRACT: Failed Back Surgery Syndrome (or FBSS) refers to patients with persistent or new pain after spinal surgery for back or leg pain. Multiple factors can contribute to its onset. We studied different techniques of data mining to determine the prognosis of patients with FBSS. Different machine learning algorithms are tested to find the best algorithm that predicts the factors that influence FBSS from the set of 305 patients operated for lumbar disc herniation. Since the data is unbalanced other criteria rather than accuracy is used as the evaluation criteria. The tools that are developed using WEKA API and the approach of feature selection test different machine learning algorithms to evaluate the best algorithm that maximize AUC or F-Measure, but on the same time maintaining false negative low. The results of the experiments are discussed and the factors that mostly influence on this syndrome are identified.

KEYWORDS: AUC, F-Measure, data mining, failed back surgery syndrome, prognosis.

I. INTRODUCTION

Medical prognosis is a prediction of the future course and outcome of a disease. Failed Back Surgery Syndrome (FBSS) is a chronic back and/or leg pain that persists after back (spinal) surgery. It is defined as "persistent or recurrent pain, mainly in the region of the lower back and legs, even after technically, anatomically successful lumbosacral spine surgeries"[1]. It is a widespread public health problem which has a considerable impact on the patient, health care system and the society. Multiple factors can contribute to the onset or development of FBSS.

Data mining classification techniques may be useful for medical prognosis and decision support in a clinical setting.Often clinical decisions are made based on doctors' intuition and experience rather than on the knowledge of rich data hidden in the database[2]. This practice leads to errors and excessive medical costs.

The aim of this article is to explore different data mining techniques in order to determine the one that best predicts the factors that influence on the FBSS. The problem of imbalanced data is raised as well as the data that we have explored is imbalanced. The class imbalance problem is a relatively new challenge since it deceptively affects classification performance. The problem of imbalanced data arises when the class distribution is too skewed[3]. Learning algorithms that do not consider class-imbalance tend to be overwhelmed by the major class and ignore the minor one [4].

II. RELATED WORK

There have been few studies on the prediction of the factors that influence FBSS. The study objective of Rodrigues F. F. et al. [5] was to report the cases of FBSS and surgical and nonsurgical etiologies based on the retrospective study of patients submitted to spinal surgery and followed in the Instituto of NeurologiaDeolindoCouto. In the study from Qidwai U. et al.[6] a Standard Fuzzy Inference System was developed around mapping the physicians' heuristics that can be used by neurosurgeons and orthopedic surgeons to predict patients' health after an operative procedure on the vertebral column just by analyzing the preoperative patient data.De Santis A.et al. in their study [7] showed that at short term follow up after lumbar disc herniation approximately 5% of the operated cases referred pain reported through Visual Analogical Scale (VAS) equal or superior to 4, which the authors considered as FBSS. Those cases were immediately selected for pain treatment at first postoperative check at one month after operation. A randomized selection was done and the FBSS patients fell in one of three categories: physiotherapy; tricyclics or both. The patients with VAS 4 or up continued to improve. The long term follow up of these patients was done and reported in PhD thesis of Alimehmeti R. [8] that permitted to identify the factors that may influence the postoperative result as well as the clinical course of the FBSS patients in one year from surgery. The data of these patients has been utilized for this study.To our knowledge

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and CommunicationEngineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

there are no studies for discovering influencing factors of FBSS with the factors and data mining approaches that we have used.

III. THE CLINICAL DATA

For the experiments the authors used the data from patients operated consecutively in two centres: 258 at the Neurosurgical Operative Unit, “Galeazzi” Institute in Milan and 47 patients at the Neurosurgical Service of “Mother Theresa” Hospital of Tirana.

The criteria for inclusion in the study were universal for all patients:

1. Patients operated for lumbar disc herniation.

2.Pain after operation not due to recurrence, or residual disc herniation.

3.Exclusion criteria: (only the patients enrolled for the group of medical treatment to whom administration of carbamazepine/clomipramine was contraindicated (patients with prostate hypertrophy, narrow angle glaucoma, cardiac arrhythmias due to incompatibility with tricyclics administration).

After cleaning the attributes that were not relevant for the study the data records were organized in 18 attributes, which represent the factors that the specialists of the field believe that mostly influence the FBSS.

In the table 1 are presented the attributes of the records, their description and values.

TABLE I
ATTRIBUTES OF THE DATA, THEIR DESCRIPTIONS AND VALUES

Number	Attribute Name	Description	Values
1.	Level	Level of the intervertebral disc which has herniated.	L5-S1,L4-L5,L3-L4, L2-L3,L1-L2
2.	Side	Side of the radicle involved (right DX; left SX).	DX,SX
3.	Age	Age of the patient in years	Numeric Min. 17, Max. 86 Mean 46.8 Standard Deviation 12.8
4.	Sex	Sex of the patient (m – male; f - female)	m,f
5.	Progressive Therapy	Steroid and non-steroid anti-inflammatory drugs administered before surgery	NSAID, Steroids, NSAID and steroids, NSAID and Physical therapy, None.
6.	Smoke	Patient’s habit	no, yes, ex
7.	Time first pain to surgery	Time from onset of symptoms to surgery	Numeric Min 1, Max. 48 Mean 7.5 Standard Deviation 9.4
8.	Previous episodes of pain	Any similar episodes of pain in the same distribution before the final that made the patient undergo operation	no, yes
9.	Preoperative VAS	Intensity of pain before surgery	0,1,2,3,4,5,6,7,8,9,10
10.	Postoperative VAS at 1 month	Intensity of pain 1 month after surgery	0,1,2,3,4,5,6,7,8,9,10
11.	Postoperative VAS at 1 year	Intensity of pain 1 year after surgery	0,1,2,3,4,5,6,7,8,9,10
12.	Lasegue grading	Angle of possible flexion of the extended lower limb	no, 10.0, 15.0, 20.0, 30.0, 35.0, 40.0, 45.0, 50.0,55.0, 60.0, 70.0 , 75.0 ,80.0
13.	Hypo or absent OTR	Attenuated or absent osteo-tendinous reflex	no, yes
14.	Motor deficit	Lowered strength of the interested muscles	no, yes
15.	Controls Sphincters	Capacity to control urinary and anal sphincters	no, yes
16.	Sensitive deficit	Alteration of the sensibility	no, yes
17.	Paravertebral contracture	Involuntary contracture of the scaffolding muscles of the vertebral column	no, yes

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and CommunicationEngineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

18.	Type of radiological examination	Radiological examinations done before surgery	MR,CT,CT and MR,CT dynamic X-ray.

(NSAID: non-steroid anti-inflammatory drugs; MR: magnetic resonance; CT: computerized tomography; hypo: under; OTR: osteotendinous reflexes; DX: right; SX: left)

The VAS was measured for each patient before the operation, one month, six months and one year after operation. The value of VAS is from 0 to 10. A value of 4 or greater at one month after operation is the criterion of having the FBSS syndrome. From 305 cases only 22 resulted to have developed the FBSS syndrome. Our data is a typical example of unbalanced data.

IV. METHODS

Data mining techniques is a powerful and important tools for the medical decision support. Recent researches have shown that application of data mining in clinical medicine has grown[9]. There are many data mining techniques each of them serves a different purpose depending on the modeling objective. These techniques have been used with a J48 decision tree algorithm [10] and with a Bayesian network [11] to discover interesting prediction rules from the data of the operated patients who underwent surgery on peripheral nerves. The two most common modeling objectives used in medicine are classification and prediction. Classification models predict categorical labels and prediction models predict continuous-valued functions. Since we will use the binary decision (weather the patient has FBSS or not) we will use classification techniques. WEKA (Waikato Environment for Knowledge Analysis), a popular open source data mining suite of machine learning algorithms, has been used to explore different classification algorithms on the data. The algorithms implemented in Weka and chosen to be applied on our data are: ADTree, BFTree, DecisionStump, FT, J48, J48graft, LADTree, LMT, NBTree, RandomForest, RandomTree, REPTree, SimpleCart, NaiveBayes, SMO, ZeroR, MultilayerPerceptron.

Most of the chosen algorithms are tree algorithms. Tree algorithms offer a supervised approach to classification. They are attractive because they provide a symbolic representation that allows easy interpretation by the users, even by non-technical people (developing patterns easily understood by physicians, both in flowchart and in text form). They are able to handle both categorical and numerical medical data and can be easily converted to a set “if-then” rules. ZeroR algorithm is a naive algorithm that predicts the majority class in the training data used as a baseline. It will serve as reference point for evaluating the performance of the classifiers considered, in the sense that one algorithm should never go below the value found with this algorithm.

To construct a predictive model we followed the four step procedure that includes familiarization with the data, data processing, modelling and evaluation of performance of the model.

V. EXPERIMENTS

The criterion correctly classified (accuracy) is frequently assumed to be the most frequently used indicator of performance on medical data. Therefore we will use this criterion to evaluate our data. Another binary attribute (failedBack) is added to the dataset which has the value 0 when the patient doesn’t have this syndrome (postoperative VAS is less than 4) and 1 when the patient has this syndrome (postoperative VAS after one month is equal or more than 4). The attribute VASPostop1Mese has been removed.

Every algorithm mentioned in the section IV has been repeated 10 times for every method and randomizing the data in every repetition.To randomize the data and to memorize the media of the criterions, one program has been developed in java using Weka API.This program uses as the method of evaluation 10 fold cross validation, hold out method 50%, 66% e 85%. The results of the method 10 fold cross validation and hold out 66% method are presented graphically in the figures 1 and 2.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and CommunicationEngineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

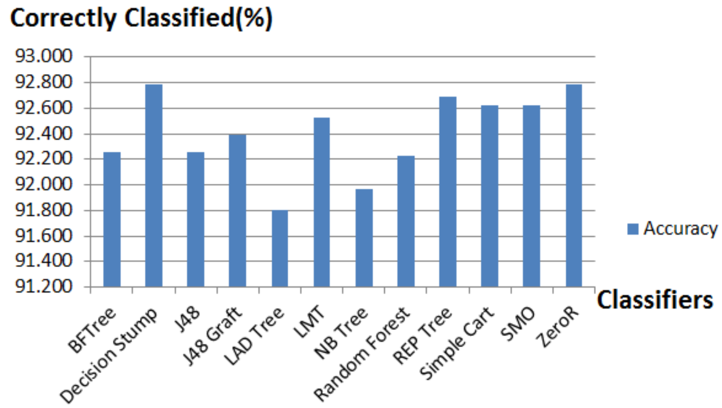

Fig. 1 The accuracy of classifiers when 10 fold cross validation is used

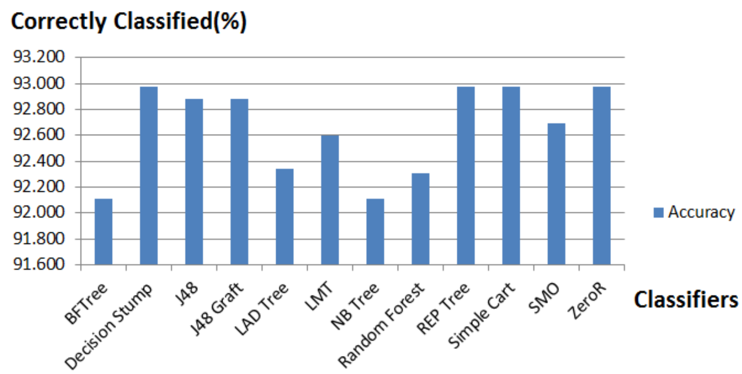

Fig. 2 The accuracy of classifiers when the method of hold out 66% is used

The results show that the greatest accuracy when the method of 10 fold cross validation and the method of hold out 66%is reached by the algorithm ZeroR followed by Tree algorithms (J48, J48graft, LMT, REP Tree e Simple Cart), DecisionStamps e SMO.The algorithm that has the highest accuracy is ZeroR even when the method of holt out 50 % and 85 % is used.This result means that the accuracy is not statistically significant. ZeroR is one of the primitive classifiers. According to the WEKA development team, this classifier predicts the majority class in the training data for all rows of test data if the class is categorical[12]. The reason why we used ZeroR is that it can be useful for determining a baseline performance as a benchmark for other learning schemes. It means that the other classifiers to be significant should have a higher accuracy than ZeroR. The reason for this result is that our data are unbalanced because negative FBSS cases are 283, compared with 22 positive FBSS cases, 92% of cases are on the same class. The class imbalance problem typically occurs when in aclassification problem, there are many more instances of some classes than others. In such cases, standard classifiers tend to be overwhelmed by the large classes and ignorethe small ones [13]. Recently, the class imbalance problem has received a lot of attention in the MachineLearning community by virtue of the fact that the performance of the algorithms used degradessignificantly if the data set is imbalanced [14].
We should use other criteria rather than accuracy to evaluate the algorithms. In the cases of unbalanced distribution, it is preferable that a classifier performs well on the minority class. The researchers use statistics such as F-Measure [15] and the area under the receiver operating characteristic curve (AUC) [16] to better evaluate the minority class. AUC and F-measure are two of the statistics which are commonly used to evaluate classifiers focusing on the importance of the

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and CommunicationEngineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

minority class [17]. F-Measure explicitly examines the performance of the classifier on the minority class.In particular, the area under the ROC curve (AUC) was shown to be a better performance criterion than the accuracy of the classification [18] and the higher the AUC, the better the classifier.

We are going to use these criteria for evaluating the selected classifiers and the technique of feature selection. This technique is often used with unbalanced data. Feature selection task involves extracting a subset of features from a given dataset by eliminating features with less or no predictive information. This technique is often used with medical imbalanced data [19]. Another program using Weka API has been created that evaluate the classifiers using F-Measure and AUC as evaluation criteria and using feature selection method. A filter created by Weka API has been used to remove the other attributes and to create groups by one, two until 17 attributes. The selected classifiers have been applied on the data having the attributes chosen from a combination of the attributes and the programs register the maximum AUC and F-Measure found for each group. The program registers for the maximum value foundfor each group, the attributes and the classifier.

The tables below 1 and 2 represent the maximum AUC or F-Measure results found for each group.

TABLE II
THE RESULTS OF THE EVALUATION FOR EACH GROUP WITH AUC

Groups	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
AUC	0.68	0.74	0.74	0.76	0.74	0.75	0.73	0.74	0.73	0.72	0.72	0.69	0.71	0.69	0.68	0.67	0.65

TABLE III
THE RESULTS OF THE EVALUATION FOR EACH GROUP WITH F-MEASURE

Groups	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
F-Measure	0.89	0.91	0.91	0.91	0.92	0.92	0.92	0.91	0.91	0.91	0.91	0.91	0.91	0.90	0.90	0.90	0.89

After analysing the attributes of respective groups that maximize the AUC and F-Measure it is realized that the attributes are not the same. To determine the attributes that mostly influence on this syndrome another program is developed which find for the groups found by the second program the group that has the minimum number of false negatives. This program finds the group of three attributes: Preoperative VAS, Paravertebral Contracture, Type of Radiological examination, with AUC value of0.739 and false negative number of 0.9. The classifier is Random Forest. This is the group that predicts FBSS better than the other groups.

VI. CONCLUSION AND FUTURE WORK

This study based on the experiments from medical field showed that we must use different criteria and methods when dealing with unbalanced data.

On the other hand from the medical point of view, this study has an important impact for the clinical decisions of everyday practice in spine surgery. It predicts which are the factors that has to be considered by the surgeon before surgery in order to have better surgical result.FBSS is an important event of spine surgery with considerable social impact. It is important to understand how the attributes influence the prevalence of this syndrome in order to predict the result of spine surgery. The algorithm with the highest value of AUC, 0.739 is Random Forest and shows that the classifier predicts well.

As future studies in this interesting field of informatics for medicine, we suggest to: develop models based on other classifiers, apply other algorithms of automatic apprehension or develop new algorithms that improve prediction experiment on larger and more diversified data set.

REFERENCES.

1. Leveque JC, Villavicencio AT, Rubin L, Bulsara K, Gorecki JP(2001), 'Spinal cord stimulation for failed back surgery syndrome', Neuromodulation,Vol. 4, pp 1-9.
2. G.Subbalakshmi, K. Ramesh M.Tech, Chinna Rao M.Tech(2011), "Decision Support in Heart Disease Prediction System using Naive Bayes". Indian Journal of Computer Science and Engineering, Vol. 2, No. 2, pp.170-176.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and CommunicationEngineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 2, February 2014

3. Weiss G.M. (2004), "Mining with rarity: a unifying framework". ACM SIGKDD Explor.Newslett. 6 , pp. 7–19.
4. Chawla N. V., Japkowicz N. and Kotcz A. (2004), "Editorial: special issue on learning from imbalanced data sets". ACM SIGKDD Explorations, 6(1) , pp.1–6.
5. Rodrigues F. F., Dozza D. C. Dozza, Oliveira C. R. , Castro R. G.(2006) , "Failed back surgery syndrome: casuistic and etiology". Arquivos de Neuro-Psiquiatria Vol. 64 (3 B), pp. 757-761.
6. Qidwai U., Shamim Sh., Enam A.(2010), "Fuzzy prediction for failed back surgery syndrome" . Applied Artificial Intelligence 24:10, pp. 881–895.
7. De Santis A., Menghetti C., Alimehmeti R. et al(2007), "Prospective randomized study for prevalence, prevention and cure of Failed Back Surgery". 55th National Congress of Italian Society of Neurosurgery, Rome.
8. Alimehmeti, R. (2007) "Studio prospetticorandomizzatomulticentricosullaprevalenza e possibileprevenzionedella failed back surgery syndrome" , PHD thesis, Milano.
9. Bellazzi R., Zupan B.(2008),"Predictive data mining in clinical medicine: Current issues and guidelines", International Journal of Medical Informatics, Vol.77, pp. 81 – 97.
10. A. Kika , I. Ninka, B. Çiço, R. Alimehmeti: Data Mining in Prediction of Surgical Outcome. In: 4th Balcan Conference in Informatics, pp 150--153.A.T.E.I Thessalonikis Publishing House (September 2009)
11. A.Kika : Using bayesian networks to discover knowledge from medical data. Proceedings. Scientific Conference: The role of the ICT in the development of Albanian Society Tiranë, 2010, pp. 287-294.
12. R.Sujatha, D.Ezhilmaran(2013),"Evaluation of Classifiers to Enhance Model Selection". International Journal of Computer Science & Engineering Technology (IJCSET),Vol.4, Issue1, pp. 16-21.
13. Chawla N. V., N. Japkowicz, Kotcz A., Editorial: special issue on learning from imbalanced data sets, in: Proceedings of the SIGKDD Explorations Newsletter, vol. 6, ACM SIGKDD Explorations, 2004, pp. 1–6
14. Japkowicz N. and Stephen, S.(2002), "The Class Imbalance Problem: A Systematic Study" , Intelligent Data Analysis, Vol. 6, Nr. 5, pp. 429-450.
15. Estabrooks, Andrew; Jo, Taeho; Japkowicz, Nathalie :A multiple resampling method for learning from imbalanced data sets, Computational Intelligence Vol. 20, Issue 1, pp: 18–36. (February , 2004).
16. Japkowicz, N. (2000). Learning from Imbalanced Data sets: A Comparison of Various Strategies. In Proceedings of the AAA1'2000 Workshop on Learning from Imbalanced Data Sets, Austin, TX.
17. Mina Alibeigi, SattarHashemi, Ali Hamzeh(2011), "Unsupervised Feature Selection Based on the Distribution of Features Attributed to Imbalanced Data Sets", International Journal of Artificial Intelligence and Expert Systems (IJAE), Vol.2, Issue 1, pp.14-22.
18. F. Provost and T. Fawcett. Robust classification for imprecise environments. Machine Learning, 42(3):203–231, 2001.
19. Wang K.J., Bunjira M. and Wang K.-M.(2013), "An improved survivability prognosis of breast cancer by using sampling and feature selection technique to solve imbalanced patient classification data", BMC Medical Informatics and Decision Making , Vol. 13, Article nr. 124.

BIOGRAPHY

Alda KIKA graduated from the University of Tirana, Faculty of Natural Sciences, Department of Computer Science in 2001. She holds a PhD diploma in Informatics. She has been teaching since her graduation in many subjects in bachelor and master courses. She is the author of several national and international journal articles and conferences in the field of data mining, medical informatics and information systems. Her work focuses on the analysis of the data using machine learning algorithms in order to discover and manage knowledge from various real application domains.

Ridvan ALIMEHMETI graduated from the University of Tirana, Faculty of Medicine in 1993. From the University of Milan in Italy he got the title of Specialist in Neurosurgery in 2001; Doctor in Medicine and Surgery in 2002; and PhD Degree in Neurological Sciences and Pain, in 2007. He successfully passed the written Exam of EANS Board in 2006 (Bratislava) after completing EANS training courses.He is neurosurgeon in charge of peripheral nerve surgery and pain treatment at Neurosurgical Department of University Hospital Center "Mother Theresa" and lecturer of Microsurgery at University of Medicine in Tirana.