Knowledge-Based Systems 73 (2015) 298–310

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Developing new machine learning ensembles for quality spine diagnosis

Indrajit Mandal *

Computer Science and Engineering Department, Rajiv Gandhi Institute of Technology, Bangalore 560 032, India

ARTICLE INFO

Article history:
Received 19 March 2012
Received in revised form 6 October 2014
Accepted 13 October 2014
Available online 19 October 2014

Keywords:
Machine learning ensembles
Multimethod assessment process
Statistical process control
Spine diagnosis
Bayesian network
Software reliability

ABSTRACT

This research work adduces new hybrid machine learning ensembles for improving the performance of a computer aided diagnosis system integrated with multimethod assessment process and statistical process control, used for the spine diagnosis based on noninvasive panoramic radiographs. Novel methods are proposed for enhanced accurate classification. All the computations are performed considering steep error tolerance rate with statistical significance level of 5% as well as 1% and established the results with corrected ***t***-tests. The kernel density estimator has been implemented to distinguish the affected patients against healthy ones. A new ensemble consisting of Bayesian network optimized by Tabu search algorithm as a classifier and Haar wavelets as the projection filter is used for relevant feature selection and attribute's ranking. The performance analysis of each method along with major findings is discussed using various evaluation metrics and concludes with propitious results. The results are compared to the existing SINPATCO platform that uses MLP, GRNN, and SVM. The optimization of machine learning algorithms is obtained using Design of Experiments scheme to achieve superior prediction accuracy. The highest classification accuracy obtained is 96.55% with sensitivity, specificity of 0.966 and 0.987 respectively. The objective is to enhance the software reliability and quality of spine disorder diagnosis using medical diagnostic system and reinforce the viability of precise treatment.

1. Introduction

A medical diagnostic system [1] generally consists of a knowledge base and some methods for solving an intended problem. On the basis of the query requested to the system, it outputs appropriate solutions and provides assistance to medical physicians in diagnosing patients. The knowledge base of such medical systems relies upon on inputs that spring up from the clinical experience of field experts [2]. Knowledge acquisition is the process to transform human expertise knowledge and skills acquired through clinical practice to software, besides being quite time consuming and labor intensive task. Common methods like Case Based Reasoning (CBR) solves the knowledge acquisition problem to some extent, in which the previous clinical cases consisting of patient's health symptoms in the form of the database is maintained with their possible remedies, past clinical decisions, preventive measures and expected diagnostic outcome measures. During patient diagnosis, the clinical database is explored for the most likely analogous past patient's record and provides the suitable diagnosis decisions for a new subject [3]. The opinion of physicians regarding the use of inference system is also appreciable in the literature [4]. The advantages of machine learning (ML) methods are that it uses mathematical models, heuristic learning and rules acquisition process for decision making and thus provides controllability, observability, stability and easily updateable by adding a new patient's record [5–7].

The application of machine learning models in the field of medicine for human disease diagnosis aids medical experts in the identification of diseases based on the symptoms at an early stage, though some diseases exhibit similar symptoms. Current advances in research for Spondylolisthesis and Disk Hernia treatment [8] since the past decades suggest that the related risk factors are identified. The state-of-the-art of technology is needed to be improved for significant accurate diagnostic decisions. The primary motivation of the research work is to design a reliable inference system and provide guidelines to caution against the selection of inaccurate inference system design for the clinical assessment of spine disorders, besides improving the quality of the diagnosis, prognosis, reduce delay in treatment and diagnosis cost [9].

2. Related works

This section provides a bird's eye view on the spine related advancements in the recent years. Kolta et al. [10] proposed a technique using the whole spine imaging based on standard DXA device

* Tel.: +91 8438366442.
E-mail address: indrajit1729@gmail.com

http://dx.doi.org/10.1016/j.knosys.2014.10.012
0950-7051/

(3D-XA) for 3D reconstruction of vertebral bodies in women. Masharawi [11] evaluated the shape characteristics of the Lumbar vertebral body (VB) and neural arch in Spondylolysis using a 3D digitizer and other parameters. Szeyko et al. [12] reported that vertebral infections caused by Coccidioides species need multidisciplinary management that always demands medical treatment and surgical intervention for stabilization. Dall'Ara et al. [13] reported that quantitative computed tomography based finite element modeling provide a quantitative and significant improvement in vertebral strength as compared to simulated dual energy X-ray absorptiometry. Yang et al. [14] introduced non-invasive fluoroscopy based image guided surgery that reduces radiation exposure due to vertebroplasty for vertebral compression fractures. Sapin-De et al. [15] gave subject-specific finite element models using low dose bi-planar X-ray devices for prediction of vertebral strength. Roberts et al. [16] achieved good accuracy for vertebral shape from lumbar radiograph using semi-automated active appearance models. Oh et al. [17] showed novel use of the liquid embolic agent Onyx injected within pseudoaneurysm, resulting in proximal and distal parent vessel closure. Wustrack et al. [18] introduced predictors of new vertebral fractures using data from the placebo arm of the HORIZON Pivotal Fractures. Roux et al. [19] showed that despite the use of PPI concomitant, there is reduced risk of new vertebral fractures during risedronate therapy as compared to placebo. Baum et al. [20] demonstrated that there is significant correlation between the volumes of vertebral bone marrow adipose tissue content to that of a volume of abdominal fat, spine bone mineral density using proton MR spectroscopy in women.

The major research contribution related to SINPATCO platform made by Rocha Neto et al. [21] in which they reported maximum accuracy of 85.9% with the implementation of SVM (linear and KMOD). In their second research article [22], they reported maximum accuracy of 90.32% by using Support Vector Machines (SVM), MLP (Multilayer Perceptron), GRNN (Generalized Regression Neural Network) and SOM (Self-Organizing Map). In this research paper, new ensembles are proposed and the results obtained over this problem using the same database are compared with the existing works. It reports higher precision, specificity and sensitivity values. Robust methods for treating vertebral column disorder including rotation forest ensemble comprising of bagging [23], Principal Components analysis (PCA) [24], decision trees [25], boosting methods [26], Haar wavelets [27], Naïve Bayes [28,29] and other machine learning algorithms are presented. A new ensemble for relevant feature selection is proposed that consists of rotation forest ensemble that employs Bayesian network [30] and Haar wavelets. The attribute ranking is procured using Ranker search algorithm. All the machine learning models are evaluated using statistical significance level of 5% and 1% along with corrected ***t***-tests [31].

This research article is succinctly organized into various sections as follows. The radiograph pelvic spine data set with its various attributes and the proposed hybrid machine learning algorithms are presented in Section 3. The theory and calculations are briefly illustrated in Section 4. The design of the proposed inference system and the results obtained from the experiments are elucidated in Section 5. Finally, the Section 6 concludes with the interpretations of the accomplished results and their significance for further research.

3. Material and methods

3.1. Data

The Vertebral Column dataset consists of 310 instances and six attributes, having one class without any missing values. There are three different categories comprising of Spondylolisthesis, Disk Hernia and normal cases. The two categories Disk Hernia and Spondylolisthesis are combined into one category labeled as "abnormal". Each patient data was collected from a medical residence period in spine surgery at The Group of Applied Research in Orthopaedics (GARO), The Centre medico-Chirurgical de Réadaptation des Massues, Lyon, France by Dr. Henry Mota (Orthopedic Surgeons' Hospital Monte Klinikum). The database is obtained from sagittal panoramic radiographs of the spine of format 30×90 cm. The data are obtained from radiographs of patients operated for herniated discs (60 patients) and Spondylolisthesis (150 subjects) and the rest 100 subjects are volunteers who do not have any conditions in the column, called normal category. Each patient is described by six biomechanical attributes which are derived from the shape and orientation of the pelvis and lumbar spine: pelvic tilt, grade of spondylolisthesis, sacral slope, lumbar lordosis angle, pelvic incidence, and pelvic radius. All these biomechanical parameters of the vertebral column are discussed elaborately in [21,22].

The radiograph's measured features along with their corresponding values are presented using parallel coordinates plot as shown in Fig. 1 after preprocessing for its visualization of high-dimensional geometry and analyzing the multivariate pelvic spine data. For showing a set of distinct points in multidimensional space, a backdrop is projected that consists of n parallel lines, projecting vertically and equally spaced. A point in the multidimensional space is reflected by a polyline with vertices lying on the parallel axes in such a way that the position of vertices on the kth axis corresponds to the kth coordinate of that point.

The machine learning ensembles have been trained with 80% of the entire clinical radiograph dataset and subsequently the degree of learning of these ensembles is tested on the remaining dataset, such that the testing is not performed on the same instances as it leads to biased evaluation. The data set is standardized in such a way that the overall mean and standard deviation is equal to 0 and 1 respectively. Then data cleaning methods are applied for dimension reduction as well as outlier's [32] removal. A new ensemble is developed for the appropriate feature selection and their ranking for the model's input and also the outliers from the dataset are removed based on the quantile information obtained statistically beyond 5% and 95%.

The research work primarily concentrates on the problem of finding optimal features needed for training and subsequent extirpate the frivolous features from the clinical dataset and additionally eliminates the latent outliers by applying felicitous techniques to develop a superior classifier for accurate prediction. In the literature, feature reduction techniques are commonly branched as extraction and selection techniques. The former converts the original features into a new feature space; the latter retains only minimum and best features from the available original ones. Both these techniques are used widely in data engineering applications for achieving computational complexity advantages while training the ensemble models. The Spearman's correlation matrix is shown in Table 1 and it reflects the degree of association among the variables of pelvic spine.

Here multiple threads are used for the generation of CHAID decision trees as it gives a simple and better interpretable method for multivariate analysis of pelvic spine data, shown in Fig. 3. The major advantages are ease of coding with modern programming languages, threads share information more efficiently, enables parallel execution and saves execution time. It's highly suitable for analyzing large medical data sets. The computation for goodness of split of the candidate attribute is optimized using a metric. The proposed pseudo code for splitting a node in the multithreaded CHAID decision tree is shown in Fig. 4. The tree is constructed using training data and then tested to obtain the LIFT-ROC curve