European Journal of Molecular &Clinical Medicine
ISSN2515-8260 Volume10, Issue 02,2023

DIAGNOSIS OF SPORTS INJURIES WITH AI MACHINE LEARNING: EXPLANATION OF INDUCED DECISIONS

Dr. S.Naveen Kumar
Associate Professor, Computer Science Engineering,
Audisankara College of Engineering and Technology (AUTONOMOUS), Gudur, Tirupati(DT),AP, India
K.Nandhini, K.Thanmayi, G.subramanyam, CH.Siva kumar.
Students, Department of Computer Science Engineering,
Audisankara College of Engineering and Technology (AUTONOMOUS), Gudur, Tirupati(DT), AP, India.

***Abstract-Medical database data can be mined* for knowledge using machine learning techniques. Various machine learning methods were employed in our application to extract diagnostic information to support judgments regarding the diagnosis of sports injuries.**

1.INTRODUCTION

cases increase every day, data on accurate diagnoses are frequently available. Similar data collection is carried out in the regular activities of specialized medical doctors. The development of diagnostic and prognostic rules as well as the resolution ofcomplex diagnostic and prognostic problems are both ideally suited for machine learning technologies. In archives of specialized hospitals and clinics, where the number of archived.

The Ljubljana University Medical Hospital's Center for Sports Medicine also collects daily records of patients who have had sports injuries. The scope of this work is the data analysis of patient records involving handball and athletic injury. Our research aims to provide systematic computer-assisted data collection and storage, intelligent analysis of stored data, support fordiagnostic judgments, and transfer of expert diagnostic knowledge from experienced specialists to young, inexperienced medical physicians. This study's goal is to clarify the murky relationships between individual anamnestic and clinical characteristics and individual diagnoses. Additionally, in order to support diagnostic decisions, a reasonable level of diagnostic accuracy must be attained, and the proposed decisions.

Numerous alternative machine learning techniques have been created in recent years. Three categories can be used to categorize machine learning techniques: inductive learning of symbolic rules (such as induction of if-then rules, decision trees, or logic programs); statistical or pattern recognition techniques (such as k-nearest neighbors or instance-based learning, discriminating Artificial neural networks (such networks with Bayesian classifiers and regression analysis), as well as Hopfield's associative memory, Kohonen's self-organizing network, and memory).In this paper, we have a bias in favour of explanation-giving systems. We have also restricted the selection of systems to various variations of top-down decision tree learners and to many variants of the Bayesian classifier that have shown to be well-suited for assisting diagnostic decision making in a variety of medical domains [Kononenko,19931.

The challenge of diagnosing sports injuries is discussed in the paper, along with the tests that were conducted, their findings, and a medical professional's assessment of the findings. How to deliver the induced information to physicians in a straightforward manner is one of the significant concerns covered in this research. In order to improve the system's capacity for an explanation, we created a general expert system shell that uses different Assistant algorithm iterations. as well as the Bayesian classifier), which offers a number of techniques permitting a visual depiction of the induced knowledge: online browsing of pie charts, tables, and decision trees with numerous types of data, such as data gains of characteristics for a suggested diagnosis.

2.DIAGNOSTIV PROBLEM

Records of patients with sports injuries are compiled daily at the Ljubljana University Medical Hospital's Center for Sports Medicine. During a patient's initial visit to the center, the injury is diagnosed and a course of therapy is advised. Patients are often treated with a variety of therapy techniques over the course of several visits to the Center, in addition to suggested at-home care and exercise.

The 118 patient records make up the current sports and handball injury database. The values of 49 characteristics to describe. Experts believe that the most crucial diagnosis characteristics such as the location of the injury, and thetest for forced movement, Despite the fact that it is obvious that different diagnostic features have various diagnostic weights. 30 diagnostic classifications (the original number) are used to classify diagnoses. Database handles over 50 diagnoses). The injury is the most prevalent diagnosis. Of ligament insertions (16% of patients had this diagnosis), making this class the most prevalent.is not much higher than other classes, which have 11% (muscle damage to the skeletal muscles), the rear side of the tight), and 10% (ankle joint injury). However, 4 classes have two training examples each, while 11 diagnostic classes are represented by a single training instance. The issue of classes with insufficient examples was only partially resolved by a suitable grouping of comparable illnesses. For instance, the diagnoses “distensions of muscles semitendinosus” and “distension of biceps femoris” were combined into the diagnostic category “injury of muscles of the back side of the tigh”; this was done because the two injuries share the same physiological cause and are both caused by injuries to muscles that are located on the back side of the thigh. Expert-defined rules in the form of training examples themselves were utilized as pre-classifiers or as generators of extra training instances for the diagnosis indicated by too few cases.

3.MACHINE LEARNING SYSTEMS AND THEIR EXPLANATION CAPABILITY

It is critical for the system to be able to defend its choices when making a diagnosis for a new patient in medicine. The user needs extensive reasoning and explanation, particularly when presented with an unexpected solution to a new problem. Only systems that offer decision explanations were employed in this study. We employed a variety of Bayesian classifier iterations as well as decision tree learners.

Decision tree learners. Assistant-R, Assistant-I, and Assistant-R2 are three variations of the Assistant algorithm that add various additions to the original Assistant method for top-down induction of decision trees [Cestnik et al., 19871]. Their strategies for attribute selection account for the majority of the differences between the algorithms: While Assistant-R and Assistant-R2 employ the algorithm ReliefF, Assistant-I uses informativity [Kononenko and Simec, 19951. In a variation of Assistant-R called Assistant-R2, decision trees are created separately for each class (diagnostic) and then combined to create a classifier for the whole domain. As opposed to Assistant-I and Assistant-R, which construct a single generic decision tree for the whole domain. Induced decision trees are fairly simple to understand and can be used to support diagnosing without using a computer. This is especially useful in situations that call for prompt decisions and in circumstances where computer interaction is psychologically unacceptable.

Decision tree learners are known to frequently provide an adequate explanation. The topmost qualities in the tree, in particular, frequently correspond to the knowledge of a domain expert. However, these approaches involve pruning [Cestnik, et al., 19871] which significantly decreases tree sizes in order to obtain generic rules. As a result, the routes between the root and the leaves are shorter and carry only a few of the most useful qualities.

Often, doctors disapprove of such trees since they consider too few characteristics and the tree Too little is known about the patients to allow for trustworthy judgments. Yet another issue is the decision tree variability - typically, a little change in the dataset results in a significant decision-tree restructuring: This further undermines doctors' confidence in the planned bothin its diagnosis and justification.

Bayesian classifiers. The naive Bayesian classifier and the semi-naive Bayesian classifier are two variations of the Bayesian classifier that have extensions for handling continuous characteristics. A development of the naive Bayesian classifier, the semi-naive Bayesian classifier explicitly looks for relationships between the values of various variables [Kononenko, 19931. Pre-discretization of continuous characteristics is necessary for both methods. The issue with (strict) pre-discretization is that minute adjustments to the continuous attribute values or bounds might have a significant impact on the probability distribution and, consequently, the classification. The naïve Bayesian classifier with the fuzzy discretization of continuous characteristics was also utilized to get over this issue. [Kononenko, 19931].

A feature's (an attribute value's) relative contribution to a diagnosis is shown in a table of conditional probabilities generated by Bayesian classifiers. The stated "weight" of a feature, or the information gain for each patient's feature, as well as the total information gains of all features that are in favor of or against the choice are used to explain a decision when it comes to a specific patient (diagnosis). One of the key benefits of such a choice, which appeals to doctors, is that all the information at hand is utilized to support the choice; this kind of justification appears to be "natural" for medical diagnosis and prognosis.

European Journal of Molecular &Clinical Medicine

ISSN2515-8260 Volume10, Issue 02,2023

4.EXPERIMENTS AND RESULTS

Since the ultimate test of the quality of learners is their performance in unseencases, experiments were performed on ten different random politicians of the data into 70% training and 30% testing examples. In this way, ten training sets E_i and ten testing sets T_i, I € [1...10] were generated. In addition, patrician followed the rule that the training set must contain at least half of all examples of each class. In this experiment, all the systems used the same training and testing sets. Results of the experiments in terms of classification, accuracy, and absolute information score. [Kononenko and bratko, 1991] are outlined below.

Results of experiments using decision tree learners. Table1 summarizes the results of the assistant algorithms. All three variants of assistants achieved approximately the same accuracy and (absolute) information score. (Note that the accuracy and information score are computed for pruned trees). The comparison of decision trees reveals that assistant I select substantially different attributes than the other two variants and also generates a slightly smaller decision tree, which is in truth slightly less accurate.

Table 1. The performance of the Assistant algorithms, all using the same parameter setting: m = 2, preparing = off, postponing = on. The number of leaves for Assistant-R2 is an average over 30 trees.

Classifier	Accuracy(%)		Inf. score		Leaves (#)
	Ā	**α**	**Ā**	**α**	
Assistant	58.2	5.8	2.19	0.28	20.9
Assistant1	62.9	5.7	2.25	0.21	26.3
Assistant R2	61.7	6.2	2.22	0.06	3.2

Results of tests with Bayesian classifiers.

The naïve Bayesian classifier's classification accuracy is greatly increased when fuzzy bounds are used, as seen in Table 2. Despite the fact that there aren't many continuous qualities, severe discretization of those attributes overstates their significance. The continuous qualities, in the doctor's perspective, are not particularly relevant for classification, and when fuzzy discretization is done correctly, it effectively reduces their influence, dramatically improving classification accuracy. It turns out that using the semi-naive Bayesian classifier in this area is unsuitable. Accuracy decreases when attribute values are joined. The outcome indicates that the qualities in this domain are mostly conditionally independent.

classifier	accuracy (%)		info.score	
	Ā	α	Ā	α
Naïve Bayes - strict	59.4	4.9	1.8	0.15
Naïve Bayes - fuzzy	69.4	3	2.32	0.19
Semi-Naïve Bayes - fuzzy	59.4	4.8	1.82	0.15

Table 2. The performance of the Bayesian classifiers with parameter setting *m=2*

5.Physician's evaluation of results

The medical expert considers the classification accuracy attained by the naïve Bayesian classifier to be satisfactory and is pleased with it. Additionally, he appreciates the explanation of the judgments made by the naive Bayesian classifier since it is, in his opinion, a good representation of how doctors diagnose their patients. Additionally, he favors the naïve Bayesian classifier sinceit classifies data using all of the relevant features.

However, the decision trees are not thought to be very transparent. In reality, the physician expert thinks that the decision tree's properties are too few and that the categorization method it uses misses crucial patient data. The judgment tree produced by Assistant-I is deemed to be illogical, but Assistant- decision R's trees accurately reflect the expert physician's understanding of the most critical features and their logical relationships.

6.Summary

With the help of multiple machine learning engines, we created a generic expert system shell that can be used to extract knowledge from information contained in medical databases. The expert system shell's explanation feature enables the user to obtain a visual representation of the knowledge that has been inferred using a variety of methods of representation, such as online browsing of decision trees, tables, and pie charts of different types of information, including information gains of attributes for a suggested diagnosis. Browsing of decision trees, tables, and pie charts of different types of information, including information gains of attributes for a suggested diagnosis.

The classification accuracy and explanation capacity of the naive Bayesian classifier with the fuzzy discretization of numerical attributes were superior to other methods, according to experimental results, and were considered to be the most suitable for actual use in the diagnosis of sports injuries.

References

1.Cestnik, et al., 1987] Cestnik, B., Kononenko, 1., and Bratko, 1. (1987). ASSISTANT 86 A knowledge elicitation tool for sophisticated users. In I. Bratko and N. Lavraö, editors, Progress in Machine Learning, pages 31—45. Sigma Press, Wilmslow.

2.[Kononenko, 1993] Kononenko, I. (1993). Inductive and Bayesian learning in medicaldiagnosis. Applied Artificial Intelligence, 7:317-337.

3.[Kononenko and Bratko, 1991] Kononenko, I. and Bratko, I. (1991). Information-based evaluation criterion for classifier's performance. Machine Learning, 6(1): 67—80.

4.[Kononenko and Simec, 1995] Kononenko, I. and Simec, E. (1995). Induction ofdecision trees using RELIEF. In: G. Della Riccia, R. Kruse and R. Viertel (eds.), Proc.of ISSEK Workshop on Mathematical and statistical methods in Artificial Intelligence, (Udine, September 1994), Springer Verlag, pp. 199-220.