PERGAMON

Expert Systems with Applications 23 (2002) 367–375

Expert Systems with Applications

www.elsevier.com/locate/eswa

Combining expert knowledge and data mining in a medical diagnosis domain

Fernando Alonso[a,*], Juan P. Caraça-Valente[a], Angel L. González[b], César Montes[b]

[a]Department of Languages and Systems, Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo, s/n, 28660 Boadilla del Monte, Madrid, Spain

[b]Department of Artificial Intelligence, Facultad de Informática, Universidad Politécnica de Madrid, Campus de Montegancedo, s/n, 28660 Boadilla del Monte, Madrid, Spain

Abstract

The medical diagnosis system described here uses underlying knowledge in the isokinetic domain, obtained by combining the expertise of a physician specialised in isokinetic techniques and data mining techniques applied to a set of existing data. An isokinetic machine is basically a physical support on which patients exercise one of their joints, in this case the knee, according to different ranges of movement and at a constant speed. The data on muscle strength supplied by the machine are processed by an expert system that has built-in knowledge elicited from an expert in isokinetics. It cleans and pre-processes the data and conducts an intelligent analysis of the parameters and morphology of the isokinetic curves. Data mining methods based on the discovery of sequential patterns in time series and the fast Fourier transform, which identifies similarities and differences among exercises, were applied to the processed information to characterise injuries and discover reference patterns specific to populations. The results obtained were applied in two environments: one for the blind and another for elite athletes. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Knowledge discovery; Data mining techniques; Expert knowledge

1. Introduction

This paper shows the results of the I4 project[1] (Intelligent Interpretation of Isokinetic Information). It describes a medical diagnosis system in the field of physiotherapy and, more specifically, muscle function assessment based on isokinetic machine data, using an expert system and data mining techniques. An isokinetic machine can be described as apparatus on which patients perform strength exercises. This machine has the peculiarity of limiting the range of movement and the intensity of effort at a constant speed (which explains the term isokinetic). Data concerning the strength exerted by the patient throughout the exercise are recorded and stored in the machine so that physicians can visually analyse the results using specialised computer software.

The information supplied by an isokinetics machine has a lot of potential uses (López-Illescas, 1993): muscular diagnosis and rehabilitation, injury prevention, training evaluation and planning, etc. However, the software built into these systems, and even the isokinetic-based diagnosis techniques themselves, still have some significant handicaps that have detracted from the success of this field:

- Standard software provides only an analogical representation of the massive data flow output by these systems. The physician is left to analyse this with no further help. This is not an easy task, as it depends almost exclusively on the personal experience of the therapist.
- Novice therapists find it enormously difficult to interpret and understand the output graphs.
- Decisions are guided by the therapist's instinct, as there are no models that can be used as a reference for most of the common injuries. Moreover, the few simple models that do exist have merely been stated by experts and are not founded on rigorous data analysis. However, there is a huge amount of stored information (performed tests) that has not been analysed to improve the procedure as a whole.

Due to the above-mentioned problems, system design should combine both practitioner expertise and knowledge that can be discovered within the data. Three objectives

* Corresponding author. Tel.: +34-91-3522546; fax: +34-91-3526388. *E-mail address:* falonso@fi.upm.es (F. Alonso).

[1] I4 was developed in conjunction with the Spanish National Centre for Sports Research and Sciences and ONCE (Spanish National Organisation for the Blind).

PII: S0957-4174(02)00072-6

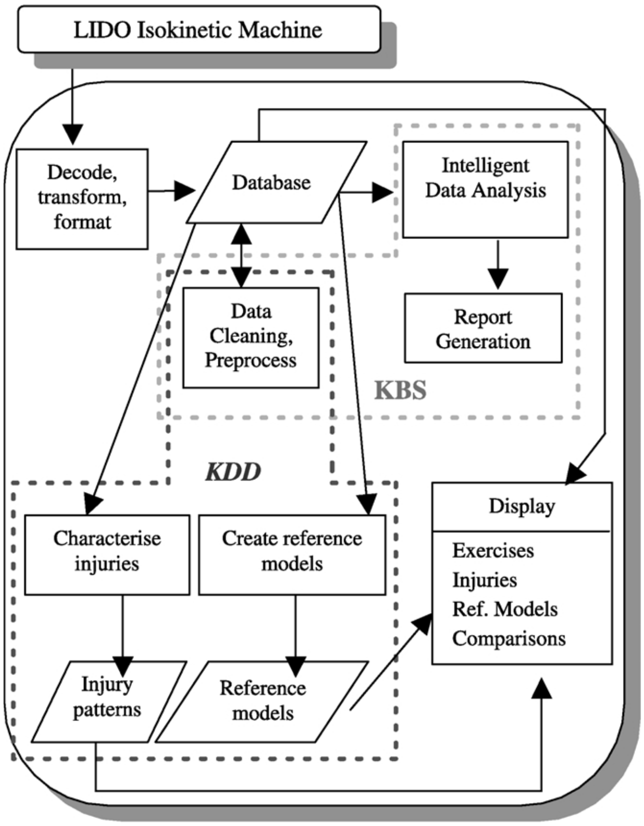

Fig. 1. I4 System architecture.

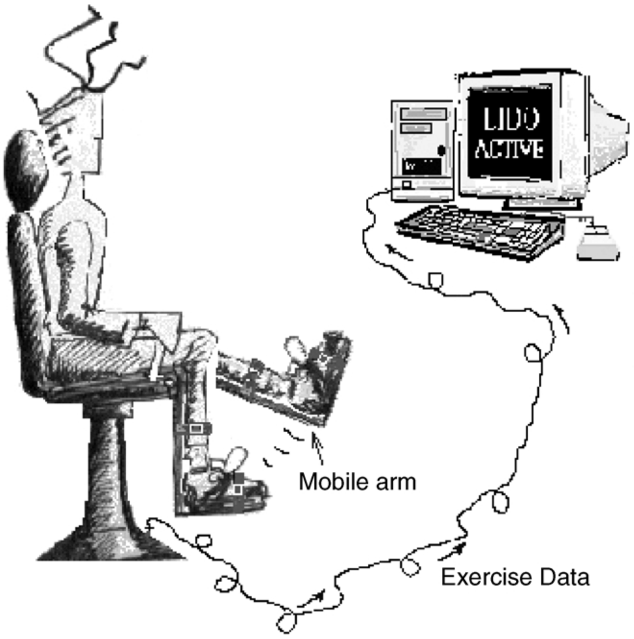

Fig. 2. Diagram of isokinetics machine use.

were therefore defined:

(a) Equip the isokinetic system with a knowledge-based system (KBS) that would perform an intelligent analysis of the strength curves output by the isokinetic test on which the assessments are based, modelling the knowledge of the expert who works the machine. This would be a valuable aid for the examining physician in detecting possible injuries.
(b) *Characterise injuries*. Bearing in mind that a huge number of isokinetic tests had already been performed and stored in a database, we aimed to find any sort of patterns useful for characterising different injuries in terms of isokinetic data. These patterns would be extremely valuable in two ways: as a useful research tool for therapists adding to the knowledge about the isokinetic shapes of common injuries and as reference models to be used for injury classification and, if possible, diagnosis. Data mining techniques based on the discovery of sequential patterns in time series were applied for this purpose.
(c) *Create reference models*. The third objective involved discovering standard patterns that characterised specific population types taken from the isokinetic data already prepared and stored in the database. For example, the process of evaluating a particular athlete against a standard curve, specific to his or her sport, is a very effective means for assessing athletes' capacity and potential for the sport they intend to go in for.

2. Overview of the system

Fig. 1 shows the architecture of the I4 system from the viewpoint of its functionalities.

The isokinetics machine includes the LIDO™ Multi-Joint II system, which supplies the data from the isokinetic tests run on patients. Seated patients move their right or left leg within a 0–90 °C flexion/extension arc (see Fig. 2). The system records the angle, as well as the strength exerted by the patient.

After the isokinetic tests have been run by the LIDO system, they are stored in the LIDO DB. As shown in Fig. 1, the first operation performed by I4 is to decode, transform and format the LIDO DB data into a standard format, correct any inaccurate or incomplete particulars and store the result in the I4 DB. This is the only I4 module that depends on the LIDO isokinetics system.

These transformed data are stored in a separate database for later processing. The Visualisation Module can display stored exercises either individually or jointly as graphs. So, this module can be used to analyse an individual exercise, to compare any pairs of exercises or even to compare an exercise with a pattern or model that is representative of a particular population group.

The Data Cleaning and Pre-processing Module automatically processes the data stored in the database in order to correct and remove any inconsistencies between the isokinetic tests. These data are processed on the basis of the